Unterrichtsintegrierte Förderung im Mathematikunterricht: Eine empirische Studie in der Primarschule

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der Universität Zürich

vorgelegt von
Meret Stöckli

Angenommen im Frühjahrssemester 2018 auf
Antrag der Promotionskommission bestehend aus
Prof. Dr. Elisabeth Moser Opitz (hauptverantwortliche Betreuungsperson)
Prof. Dr. Fritz Staub

Zürich, 2019

Zusammenfassung

Eine Förderung von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen im regulären Mathematikunterricht ist für Lehrpersonen eine grosse Herausforderung. Ausgehend von aktuellen Forschungserkenntnissen zu inhaltlichen Schwierigkeiten rechenschwacher Schülerinnen und Schüler und deren Förderung, werden zentrale Aspekte einer unterrichtsintegrierten Mathematikförderung in der dritten Klasse der Primarschule herausgearbeitet. Dazu gehören eine Fokussierung auf zentrale Inhalte der Mathematik, produktives Üben, der Einsatz von geeigneten Veranschaulichungen und Hilfsmitteln, gemeinsame Lernsituationen sowie eine adaptive Lernbegleitung. Die Frage nach der Wirksamkeit einer solchen Mathematikförderung wird mithilfe eines quasi-experimentellen Studiendesigns (N = 888) untersucht. Es werden zwei Interventionsgruppen (schriftliche Anleitungen vs. intensivere Begleitung für die Lehrpersonen) und eine Kontrollgruppe eingesetzt. Zudem wird untersucht, welche Variablen auf Individual- und Klassenebene einen Einfluss auf die Mathematikleistung haben. Die Ergebnisse zeigen, dass die unterrichtsintegrierte Mathematikförderung in einer der beiden Interventionsgruppen wirksam war. Mögliche Ursachen für die unterschiedlichen Ergebnisse bezüglich der Wirksamkeit für die beiden Interventionsgruppen werden analysiert und diskutiert. Zudem konnten auf Individual- und auf Klassenebene Variablen identifiziert werden, die die Mathematikleistung beeinflussen.

Abstract

Fostering low achievers in mathematics during regular instruction is a major challenge for teachers. Important factors for mathematical fostering in regular classroom settings in third grade of primary school are presented based on the current state of research on content related difficulties of low achievers in mathematics and on supporting these students. Such factors are: focussing on important topics, productive practicing of arithmetic, the use of suitable visualisations and manipulatives, joint learning situations and adaptive teaching. With a quasi-experimental design (N = 888) the effectiveness of such an approach to fostering low achievers in mathematics in a regular classroom setting is examined using two interventionsgroups (written instruction vs. more intense support for teachers) and a controlgroup. Additionally, variables on individual and on class level with an influence on mathematical achievement are determined. Results reveal a positive effect for the fostering approach in one interventionsgroup. Possiple reasons for the divergent results are analysed and discussed. Moreover, variables on individual and class level influencing math achievement have been identified and are presented.

Danksagung

Ein besonderes Dankeschön möchte ich Prof. Dr. E. Moser Opitz für die vielfältige fachliche und persönliche Unterstützung in den vergangenen Jahren aussprechen. Mein Dank gilt ausserdem Prof. Dr. A. Heinze, Prof. Dr. A. Lindmeier und Prof. Dr. O. Lüdtke, die mir während meines Auslandaufenthaltes am IPN in Kiel mit fachlichem Rat zur Seite standen. Zudem wäre diese Arbeit ohne die Zusammenarbeit mit meinen geschätzten Kolleginnen Dr. M. Pfister und L. Reusser nicht möglich gewesen. Ebenfalls äusserst hilfreich war die statistische Beratung durch Dr. U. Grob. Ein herzlicher Dank geht auch an die Lehrpersonen und Schülerinnen und Schüler, die an der Untersuchung teilgenommen haben sowie an alle Hilfskräfte, die bei der Datenerhebung zum Einsatz kamen. Zu guter Letzt gilt ein besonderes Dankeschön auch meinem Mann, meiner Familie sowie meinen Freundinnen und Freunden.

Inhaltsverzeichnis

1. Einleitung

Der in den vergangenen Jahren in vielen Ländern (z.B. Schweiz, Deutschland, USA, England) initiierte Wechsel von einem separativen hin zu einem integrativen Schulsystem, in dem u.a. Schülerinnen und Schüler mit Lernschwierigkeiten in Regelklassen unterrichtet werden, hat dazu geführt, dass sich die empirische Unterrichtsforschung vermehrt mit Fragen effektiver Unterrichtsgestaltung und der Förderung von Lernenden mit unterdurchschnittlichen schulischen Leistungen in integrativen Schulsettings beschäftigt (Kiel, 2017; Semper, Mende & Berkemeyer, 2017). Für das Fach Mathematik lässt sich zudem ein intensiver wissenschaftlicher Diskurs feststellen, in dem mit Verweis auf den Begriff der Rechenschwäche bzw. der Rechenstörung sowohl grundsätzliche Fragen zu Problemen von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen (z.B. inhaltliche Schwierigkeiten, Einflussfaktoren; Andersson & Östergren, 2012; Geary, Hoard, Nugent & Bailey, 2012; Vanbist, Ghesquière & de Smedt, 2014) als auch spezifische Überlegungen für eine Förderung dieser Personengruppe im Rahmen des Regelunterrichts thematisiert werden (z.B. Gaidoschik, 2016; Häsel-Weide, 2016; Scherer & Moser Opitz, 2010). Zudem gibt es auch vermehrt Forschung zu Interventionsstudien im Bereich der Förderung rechenschwacher Schülerinnen und Schüler (z.B. Ennemoser, Sinner & Krajewski, 2015; Moser Opitz et al., 2017; Wißmann, Heine, Handl & Jacobs, 2013) sowie Metastudien, die sich mit dieser Frage auseinandersetzen (z.B. Gersten et al., 2009; Ise, Dolle, Pixner & Schulte-Körne, 2012). Dabei zeigt sich, dass Ansätze, die eine verstehensorientierte Förderung bei Rechenschwäche umsetzen, wirksam sind. Als bedeutsame Aspekte einer solchen Förderung konnten u.a. eine inhaltliche Fokussierung auf zentrale Inhalte der Mathematik, produktives Üben, der Einsatz von Veranschaulichungen und Arbeitsmitteln sowie das Anregen zum Verbalisieren belegt werden. Dennoch verbleiben aktuell noch verschiedene offene Fragen mit Blick auf die Förderung von Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen in der Regelschule.

Diskutiert wird derzeit insbesondere, welche organisatorische Form der Förderung bzw. welche didaktischen Settings für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen im Rahmen einer integrativen Schule sinnvoll sind (z.B. Ise et al., 2012). Ein grosser Teil der bisherigen Förderansätze bei Rechenschwäche wurde als Kleingruppenförderung ausserhalb des Klassenzimmers umgesetzt (z.B. Bryant, Bryant, Gersten, Scammacca & Chavez, 2008; Ennemoser, Krajewski & Sinner, 2017; Wißmann et al., 2013). Es finden sich aber auch vereinzelt Studien, die eine Förderung untersuchen, die stärker klassenintegriert organisiert ist (z.B. Freesemann, 2014; Woodward & Brown, 2006). In diesem Zusammenhang ist zu erwähnen, dass an aktuellen Individualisierungs- und Differenzierungsbemühungen im Rahmen des Regelunterrichts kritisiert wird, dass die Gefahr einer sozialen Vereinzelung besteht (z.B. Moser Opitz, 2009). Entsprechend wird gefordert, dass Massnahmen der Individualisierung im Unterricht ergänzt werden durch gemeinsame Lernsituationen. Solche gemeinsamen Erarbeitungs- und Reflexionsprozesse werden auch aus mathematikdidaktischer und entwicklungslogischer Sicht als bedeutsam für die mathematische Entwicklung angesehen (z.B. Feuser, 2013; Scherer & Moser Opitz, 2010). In der vorliegenden Arbeit soll daher der Frage nachgegangen werden, ob eine verstehensorientierte Mathematikförderung, die integriert in den regulären Mathematikunterricht in der Primarschule umgesetzt wird und gemeinsame Erarbeitungs- und Reflexionsprozesse ermöglicht, wirksam ist.

Inhaltlich baut die Mathematikförderung, die in dieser Arbeit untersucht wird, auf dem aktuellen Forschungsstand zu mathematischen Schwierigkeiten von rechenschwachen Schülerinnen und Schülern auf. Die Mathematikförderung soll dabei auf mathematische Inhalte bzw. Kompetenzen fokussieren, die für die weitere mathematische Entwicklung als bedeutsam angesehen werden und von denen bekannt ist, dass sie rechenschwachen Schülerinnen und Schülern Mühe bereiten (z.B. dezimales Stellenwertsystem, Rechenstrategien, Mathematisieren, Zählen, Aufbau von Faktenwissen; z.B. Andersson, 2010; Geary, Hoard, Byrd-Craven & DeSoto, 2004; Jordan, Hanich & Kaplan, 2003; Moeller, Pixner, Zuber, Kaufmann & Nuerk, 2011; Moser Opitz, 2013). Um ein verstehensorientiertes Lernen zu ermöglichen, werden mathematikdidaktische Überlegungen zu geeigneten Veranschaulichungen und Arbeitsmittel, zu produktiven Übungsformaten und zur Versprachlichung mathematischer Handlungen und Denkprozesse berücksichtigt (z.B. Krauthausen & Scherer, 2007; Scherer & Moser Opitz, 2010; Van de Walle, 2007). Nebst gemeinsamen Lernsituationen im Rahmen von Einführungs- und Reflexionssequenzen werden in der zu untersuchenden Mathematikförderung zudem gezielt Massnahmen für eine Individualisierung und Differenzierung der Unterrichtssequenzen (z.B. Scherer & Moser Opitz, 2010) sowie für eine adaptive Lernbegleitung durch die Lehrperson eingesetzt (z.B. Beck et al., 2008; Parsons et al., 2018; Pfister, 2016), so dass alle Lernenden auf ihrem aktuellen Lernniveau gefördert werden und eine optimale Unterstützung sowie Strukturierung der individuellen Lernprozesse ermöglicht wird.

Um die Wirksamkeit einer solchen unterrichtsintegrierten Mathematikförderung zu untersuchen, wurde eine Interventionsstudie mit quasi-experimentellem Design durchgeführt[1]. 888

[1] Das Projekt wurde vom Schweizerischen Nationalfonds unter dem Titel „Rechenschwache Schülerinnen und Schüler unterrichtsintegriert fördern“ finanziert (Nr. 100013-134652).

Schülerinnen und Schüler aus 58 Schulklassen des dritten Primarschuljahres[2] haben verteilt auf zwei Interventionsgruppen und eine Kontrollgruppe an der Untersuchung teilgenommen. Im Rahmen der vorliegenden Arbeit soll basierend auf dieser Interventionsstudie folgenden Fragen nachgegangen werden:

- Kann eine verstehensorientierte Mathematikförderung, die auf zentralen Forschungserkenntnissen zur Förderung rechenschwacher Schülerinnen und Schüler aufbaut, auch im Rahmen des regulären Mathematikunterrichts wirksam umgesetzt werden?
- Wirkt sich eine intensivere Begleitung der Lehrpersonen durch Fortbildungstreffen im Vergleich zu einer Umsetzung, bei der nur schriftliche Anleitungen und Materialien für die Förderung zur Verfügung gestellt werden, positiv auf die Leistungen der Schülerinnen und Schüler aus?
- Welche Merkmale auf Individual- und Klassenebene haben einen Einfluss auf die Mathematikleistungen der Schülerinnen und Schüler?

Für die Auswertung der Daten werden in erster Linie Mehrebenenanalysen eingesetzt, da sich diese für hierarchisch strukturierte Daten eignen (vgl. Snijders & Bosker, 2012).

Aufgrund der Bedeutsamkeit von Forschungserkenntnissen zum Thema Rechenschwäche für die zu untersuchende unterrichtsintegrierte und verstehensorientierte Mathematikförderung, wird in dieser Arbeit in einem *ersten Schritt* eine begriffliche Klärung vorgenommen (Kap. 2). Einsteigend wird dafür zunächst auf die Definitionen von Rechenstörung im Rahmen der ICD-10 und des DSM-V eingegangen, und im Anschluss daran wird der Begriff der Rechenschwäche erläutert (Kap. 2.1). Weiter werden Modelle mathematischer Verarbeitung und Entwicklung thematisiert, um einen Überblick über zentrale mathematische Entwicklungsschritte zu erhalten und bedeutsame Inhalte und Kompetenzen für die mathematische Entwicklung zu identifizieren (Kap. 2.2). Zudem werden verschiedene Faktoren besprochen, die im Rahmen eines multikausalen Verständnisses von Rechenschwäche als bedeutsam für die Entstehung von mathematischen Schwierigkeiten angesehen werden, und daher für die Gestaltung der Mathematikförderung zu beachten sind (z.B. arbeitsgedächtnisentlastende Massnahmen; Kap. 2.3). Anschliessend wird der aktuelle Forschungsstand zu inhaltlichen Schwierigkeiten rechen-

---

[2] In der Schweiz werden die Schuljahre offiziell ab Eintritt in den Kindergarten gezählt. Demnach besuchen Kinder in der Schweiz im ersten und im zweiten Schuljahr den Kindergarten und vom dritten bis zum achten Schuljahr die Primarschule, bevor ein Wechsel in die gegliederte Oberstufe stattfindet. In dieser Arbeit wird mit Verweis auf den deutschsprachigen, internationalen Fachdiskurs jedoch von Klassenstufe gesprochen und mit der Zählung nach dem Kindergarten begonnen, sodass die in der Arbeit als dritte Klasse beschriftete Klassenstufe in der Schweiz offiziell als fünftes Schuljahr bezeichnet wird.

schwacher Schülerinnen und Schüler ausführlich vorgestellt, mit dem Ziel die zentralen mathematischen Inhalte zu bestimmen, die in einer mathematischen Förderung aufzunehmen sind (Kap. 2.4). In einem *zweiten Schritt* werden Erkenntnisse zur Förderung rechenschwacher Schülerinnen und Schüler dargestellt, so dass weitere zentrale Aspekte identifiziert werden können, die für die Gestaltung einer Mathematikförderung zu berücksichtigen sind (Kap. 3). Dazu werden empirische Erkenntnisse aus Interventionsstudien zusammengefasst, in denen Schülerinnen und Schüler mit mathematischen Schwierigkeiten gefördert wurden (Kap. 3.1). Weiter werden ausgewählte didaktische (Kap. 3.2) sowie inhaltsübergreifende und inhaltsspezifische mathematikdidaktische Überlegungen (Kap. 3.3) zur Förderung rechenschwacher Schülerinnen und Schüler ausführlich thematisiert. Zudem wird auf die Bedeutung professioneller Kompetenzen und professionellen Wissens von Lehrpersonen für die Gestaltung eines qualitativ hochstehenden Mathematikunterrichts eingegangen und es wird geklärt, was hinsichtlich Fortbildungsmassnahmen von Lehrpersonen zu beachten ist (Kap. 3.4). In einem *dritten Schritt* (Kap. 4) werden die Erkenntnisse aus den vorangehenden Kapiteln zusammenfassend dargestellt und offene Fragen zur Förderung von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen aufgezeigt (Kap. 4.1). Darauf basierend werden didaktische und mathematikdidaktische Schlussfolgerungen für die Gestaltung der in dieser Arbeit zu untersuchenden mathematischen Förderung in der dritten Klasse und deren Umsetzung im Rahmen eines regulären Mathematikunterrichts gezogen (Kap. 4.2). Zudem wird die Förderung im Rahmen des Angebots-Nutzungs-Modells (Fend, 1981; Helmke & Weinert, 1997) eingeordnet.

In einem *vierten Schritt* wird die Untersuchung, auf der die vorliegende Arbeit beruht, ausführlich vorgestellt (Kap. 5). Es werden die Fragestellungen (Kap. 5.1), das Untersuchungsdesign (Kap. 5.2) und die Intervention (Kap. 5.3) präsentiert. Weiter wird auf die verwendeten Messinstrumente eingegangen (Kap. 5.4) und die Stichprobe beschrieben (Kap. 5.5). Das Kapitel schliesst mit Ausführungen zu den verwendeten Auswertungsmethoden (Kap. 5.6). In einem *fünften Schritt* werden die Ergebnisse präsentiert (Kap. 6), wobei zuerst auf Ergebnisse der ersten Fragestellung zu Einflussfaktoren der Mathematikleistungen auf Individual- und Klassenebene eingegangen wird (Kap. 6.2) und danach die Ergebnisse der zweiten und dritten Fragestellung zur Wirksamkeit der unterrichtsintegrierten Mathematikförderung vorgestellt werden (Kap. 6.3). Zudem werden weiterführende Mehrebenenmodelle diskutiert, die zur Klärung der Ergebnisse beitragen (Kap. 6.4), sowie zusätzliche Analysen für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen präsentiert (Kap. 0). In einem *letzten Schritt* werden die Ergebnisse abschliessend zusammengefasst und diskutiert (Kap. 7). Die Arbeit endet mit Überlegungen zu den Grenzen der vorliegenden Untersuchung (Kap. 7.3) sowie einem Fazit, indem Schlussfolgerungen für zukünftige Untersuchungen gezogen werden (Kap. 7.4).

2. Rechenschwäche, Rechenstörung und Dyskalkulie

2.1 Begriffliche Klärungen

Bis heute gibt es in der Fachliteratur keine einheitliche Begriffsverwendung für Schülerinnen und Schüler, die Schwierigkeiten beim Mathematiklernen haben und unterdurchschnittliche bis stark unterdurchschnittliche mathematische Leistungen aufweisen. Im deutschsprachigen Raum findet man häufig die Begriffe Rechenstörung und Dyskalkulie, wobei diese insbesondere von Autorinnen und Autoren aus der Psychologie und der Medizin verwendet werden. In der pädagogischen Fachliteratur findet man häufiger Begriffe wie Rechenschwäche, mathematische Lernschwäche und mathematische Schulleistungsschwäche. Beispiele für Begriffe aus dem englischsprachigen Raum sowie dafür verwendete Abkürzungen in Studien sind: dyscalculia, developmental dyscalculia (DD), mathematical (learning) disabilities (MLD), mathematic (learning) difficulties (MD), arithmetic (learning) disability (z.B. Devine, Soltész, Nobes, Goswami & Szűcs, 2013), low achievement in mathematics (LA; z.B. Mazzocco, Devlin & McKenney, 2008) und special educational needs in mathematics (SEN; z.B. Kroesbergen & van Luit, 2003). Interessant ist hier, dass manchmal der Begriff Behinderung (disability) und manchmal der Begriff der Schwierigkeiten (difficulties) verwendet wird. Das zeigt, dass es unterschiedliche Perspektiven auf das Phänomen schwacher Mathematikleistungen gibt und es im wissenschaftlichen Diskurs bisher nicht gelungen ist, sich auf eine einheitliche Definition zu einigen (Murphy, Mazzocco, Hanich & Early, 2007). In diesem Kapitel werden daher gängige Definitionen vorgestellt und voneinander abgegrenzt. Ausgangspunkt sind dabei die Klassifikationssysteme ICD-10 und DSM-V. Das Kapitel schliesst mit Angaben zu Prävalenzschätzungen von Rechenstörung bzw. Rechenschwäche und einer Beschreibung der Verwendung des Begriffs der Rechenschwäche in dieser Arbeit.

2.1.1 ICD-10

Definition von Rechenstörung

Sehr häufig finden sich in der deutschsprachigen Fachliteratur der Begriff der Rechenstörung und der fast durchgehend synonym verwendete Begriff der Dyskalkulie (z.B. Hasselhorn & Schuchardt, 2006; Kuhn, Raddatz, Holling & Dobel, 2013). Dabei wird häufig auf die ICD-10 (International Statistical Classification of Diseases and Related Health Problems) der Weltgesundheitsorganisation (WHO) verwiesen (Dilling, Mombour & Schmidt, 2014). Eine Rechenstörung wird in dieser Klassifikation verstanden als umschriebene Entwicklungsstörung schulischer Fertigkeiten, die in der Kindheit beginnt (Dilling et al., 2014). Bei den Betroffenen müssen bereits grundlegende Rechenfertigkeiten beeinträchtigt sein, damit eine Rechenstörung diagnostiziert wird: „Das Defizit betrifft die Beherrschung grundlegender Rechenfertigkeiten wie Addition, Subtraktion, Multiplikation und Division, weniger die höheren mathematischen Fertigkeiten, die für Algebra, Trigonometrie, Geometrie und Differential- sowie Integralrechnungen benötigt werden“ (Dilling et al., 2014, S. 338). Gemäss den diagnostischen Kriterien der ICD-10 muss „[…] ein Wert in einem standardisierten Rechentest vor[liegen], der mindestens zwei Standardabweichungen unterhalb des Niveaus liegt, das aufgrund des chronologischen Alters und der allgemeinen Intelligenz des Kindes zu erwarten wäre“ (Dilling & Freyberger, 2014, S. 290). Somit müssen die mathematischen Leistungen eines Kindes mit einer Rechenstörung nicht nur das oben genannte Mindestmass an Diskrepanz zu den mathematischen Leistungen gleichaltriger Kinder aufweisen, sondern es muss auch eine entsprechende Diskrepanz zur allgemeinen Intelligenz des Kindes vorliegen. Dies wird als (doppeltes) Diskrepanzkriterium bezeichnet.

Ein weiterer bedeutsamer Aspekt der Definition von Rechenstörung gemäss ICD-10 ist das Verständnis, dass eine Rechenstörung eine Teilleistungsstörung ist, die sich nur auf mathematische Schwierigkeiten bezieht und deshalb von den kombinierten Störungen schulischer Fertigkeiten (F81.3) abgegrenzt wird (Dilling et al., 2014; Dilling & Freyberger, 2014). Letztere wird z.B. für Fälle verwendet, in denen zusätzlich zu einer Rechenstörung auch eine Lese-/Rechtschreibstörung vorhanden ist. Damit die Diagnose Rechenstörung gestellt werden kann, müssen somit auch die Lese- und Rechtschreibfähigkeiten geprüft werden, wobei diese für die Diagnose Rechenstörung im Normbereich liegen müssen (Dilling et al., 2014). Zudem gibt es weitere Ausschlusskriterien für die Diagnose einer Rechenstörung. So darf diese „[…] nicht allein durch eine allgemeine Intelligenzminderung oder eine eindeutig unangemessene Beschulung […]“ (Dilling et al., 2014, S. 338) bedingt sein. Unter unangemessener Beschulung wird z.B. das Fehlen eines regelmässigen Schulbesuchs und nicht ein als qualitativ nicht ganz so hochstehender Unterricht verstanden (Moser Opitz & Ramseier, 2012; Thomas, Schulte-Körne & Hasselhorn, 2015). Der Ausschluss einer allgemeinen Intelligenzminderung bedeutet, dass bei Schülerinnen und Schülern mit einem IQ < 70 nicht eine Rechenstörung, sondern eine Intelligenzminderung diagnostiziert wird (Dilling et al., 2014). Zusätzlich muss auch ausgeschlossen werden, dass die Rechenstörung auf Defizite beim Sehen und Hören zurückzuführen ist, durch eine neurologische Störung bedingt ist oder durch eine Krankheit ausgelöst wurde (Dilling et al., 2014).

Bezüglich des Grenzwerts für die Diskrepanz zwischen IQ und Mathematikleistung gibt es unterschiedliche Richtwerte (Kuhn et al., 2013). Gemäss den Leitlinien der Deutschen Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie (2007) muss die Diskrepanz zwischen Mathematikleistung und Intelligenz mindestens 1.5 Standardabweichungen betragen, wobei die Mathematikleistung bei einer durchschnittlichen Intelligenz in einem standardisierten Test den 10. Prozentrang nicht überschreiten darf. Fischbach et al. (2013) berichten jedoch, dass in der Forschung teilweise auch eine Diskrepanz von nur einer Standardabweichung sowohl mit Bezug auf die Diskrepanz der Mathematikleistung im Vergleich zu Kindern der gleichen Altersstufe als auch mit Bezug auf die Diskrepanz zwischen Mathematikleistung und Intelligenz für die Diagnose einer Rechenstörung angewendet wird, was jeweils einem Prozentrang von höchstens 16 entspricht. Vukovic und Siegel (2010) kommen mit Verweis auf diverse Studien zum Schluss, dass in der Forschung zu Rechenstörung in der Regel Cut-off-Werte für Mathematikleistungen verwendet werden, die tiefer als der 10., teilweise aber auch tiefer als der 16. Prozentrang liegen. Diese unterschiedlichen Grenzwerte werfen Fragen zur Zuverlässigkeit des Diskrepanzkriteriums auf.

Kritik am Diskrepanzkriterium

Das doppelte Diskrepanzkriterium der ICD-10 besagt, dass die Mathematikleistung nicht nur deutlich unterhalb des altersbezogenen Normbereichs liegen, sondern auch ein Mindestmass an Unterschied zwischen allgemeiner Intelligenz und der mathematischen Leistung eines Kindes vorliegen muss, damit die Diagnose Rechenstörung gestellt werden kann (Dilling et al., 2014). Das zweite Diskrepanzkriterium bezieht sich auf die verlangte Diskrepanz zwischen Mathematikleistung und allgemeiner Intelligenz und wird in der Fachliteratur häufig kritisiert (Ehlert, Schroeders & Fritz-Stratmann, 2012). Ihren Ursprung hatte diese Kritik bereits in den 1970er-Jahren, wobei sich diese damals auf die Diagnosekriterien für Legasthenie (Lese-/Rechtschreibstörung) bezog (Ehlert et al., 2012).

Eine erste Kritik bezieht sich darauf, dass empirisch eine hohe Korrelation zwischen Intelligenz und Mathematikleistung besteht und mathematische Kompetenzen in einigen Intelligenztheorien als Teilfacette der Intelligenz konzipiert sind, was dazu führt, dass entsprechende Intelligenztests Aufgaben enthalten, die eigentlich mathematische Kompetenzen erfassen (vgl. Ehlert et al., 2012). Es stellt sich daher grundsätzlich die Frage, wie sinnvoll Diagnosen sind, bei denen eines der ausschlaggebenden Kriterien ein Mindestmass an Diskrepanz zwischen zwei Konstrukten ist, die per se hoch korrelieren. Ein weiterer Kritikpunkt bezieht sich auf die Feststellung, dass bei niedrigen Intelligenzwerten nur bei extrem niedrigen Mathematikleistungen das Diskrepanzkriterium erreicht werden kann (Fischbach et al., 2013; Francis et al., 2005). Wie Fischbach et al. (2013) aufzeigen, führt die Berücksichtigung des zweiten Diskrepanzkriteriums dazu, dass insbesondere intelligentere Kinder mit leicht unterdurchschnittlichen Mathematikleistungen als rechengestört diagnostiziert werden. Daher wird häufig die Befürchtung geäussert, dass Kinder, die auf zusätzliche schulische Unterstützung sowie auf mit der Diagnose verbundene Beteiligungen an den Kosten für schulische und/oder ausserschulische Massnahmen angewiesen sind, aufgrund einer fehlenden Diskrepanz der mathematischen Leistung im Vergleich zur Intelligenz von notwendigen Hilfsmassnahmen ausgeschlossen bleiben (Ehlert et al., 2012; Moser Opitz & Ramseier, 2012; Thomas et al., 2015). Ehlert et al. (2012) belegen zudem, dass die verwendeten Berechnungsmethoden für die Erfüllung des Diskrepanzkriteriums zu unterschiedlichen Ergebnissen hinsichtlich der Diagnose führen. Ebenso wird häufig vernachlässigt, dass bei jeder psychologischen Messung Messfehler auftreten und somit eigentlich gar kein eindeutiger Grenzwert, sondern nur ein Grenzwertbereich angegeben werden kann, in dem die Leistungen eines Kindes mit grosser Wahrscheinlichkeit tatsächlich liegen (Ehlert et al., 2012). Ein starres Festhalten an angeblich eindeutigen Grenzwerten ist daher aus verschiedenen testtheoretischen Gründen nicht sinnvoll und muss für die Diagnosestellung berücksichtigt werden.

Ein weiterer zentraler Kritikpunkt am Diskrepanzkriterium bezieht sich auf die grundsätzliche Frage, ob unterschiedliche Intelligenzwerte an sich einen Einfluss auf die Ausprägung einer Rechenstörung haben. Aus pädagogischer Sicht ist diese Kritik am bedeutsamsten, da damit die Frage verbunden ist, ob von Rechenschwierigkeiten betroffene Kinder mit und ohne Erfüllung des Diskrepanzkriteriums eventuell auch auf unterschiedliche schulische Unterstützungsmassnahmen angewiesen sind (Mähler, 2019). Moser Opitz (2013) konnte belegen, dass sich bei Kindern mit niedrigen mathematischen Leistungen mit durchschnittlichem IQ (95–115) und Kindern mit niedrigen mathematischen Leistungen und zusätzlich unterdurchschnittlichem IQ (75–94) bei der Beherrschung der grundlegenden Elemente des mathematischen Basisstoffes keine signifikanten Unterschiede bezüglich Ausmass sowie Art und Weise der Schwierigkeiten zeigen. Ebenfalls keine Unterschiede in den mathematischen Leistungen zwischen Kindern mit einer Rechenstörung mit durchschnittlichem und mit unterdurchschnittlichem IQ fanden Jiménez Gonzalez und Garcia Espinel (2002), Ehlert et al. (2012) und Kuhn et al. (2013). Zudem konnten auch für verschiedene Komponenten des Arbeitsgedächtnisses keine signifikanten Unterschiede zwischen Kindern mit einer Rechenstörung, die das zweite Diskrepanzkriterium erfüllen, und solchen, die es nicht erfüllen, festgestellt werden (Maehler & Schuchardt, 2011). Die genannten Studien weisen darauf hin, dass eine Unterscheidung zwischen rechenschwachen Schülerinnen und Schülern mit durchschnittlichem und mit unterdurchschnittlichem IQ wenig sinnvoll ist und somit für die Diagnosestellung als fragwürdig bezeichnet werden kann.

An dieser Stelle sollen auch noch zwei allgemeine Kritikpunkte an der ICD-10-Definition von Rechenstörung genannt werden, die sich nicht direkt auf das Diskrepanzkriterium beziehen, aber dennoch bedeutsam sind für die vorliegende Arbeit. So wird an der Definition kritisiert, dass Kinder, die zusätzlich zu einer Rechenstörung auch eine Lese-/Rechtschreibstörung aufweisen, von der Diagnose ausgeschlossen bleiben, da Studien aufzeigen konnten, dass diese Kombination häufiger auftritt, als dies früher angenommen wurde (Dirks, Spyer, van Lieshout & de Sonneville, 2008; Schwenck & Schneider, 2003). In einer repräsentativen Schweizer Stichprobe im zweiten Primarschuljahr wiesen sogar zwei Drittel der Kinder mit einer Rechenstörung auch eine Lese-/Rechtschreibstörung auf und nur ein Drittel eine isolierte Rechenstörung (von Aster, Schweiter & Weinhold Zulauf, 2007). Ein nicht zu vernachlässigender Anteil von Kindern mit gravierenden Rechenschwierigkeiten hat somit auch Schwierigkeiten im Bereich des Lesens und Schreibens und sollte daher in der Forschung zu Rechenschwierigkeiten ebenfalls berücksichtigt werden. Es muss deshalb kritisch hinterfragt werden, dass das Vorliegen einer Lese-/Rechtschreibstörung als zentrales Ausschlusskriterium in der ICD-10 definiert wird (Moser Opitz & Ramseier, 2012).

Ein zweiter Kritikpunkt bezieht sich auf das Verständnis der Rechenstörung als andauernde und früh beginnende Entwicklungsstörung in der ICD-10. So konnte eine Langzeitstudie belegen, dass ein Drittel der Kinder, die aufgrund ihrer mathematischen Leistungen vom ersten bis zum vierten Schuljahr als persistent rechenschwach bezeichnet wurden, in einzelnen Jahren mathematische Leistungen gezeigt haben, die dennoch im Normbereich lagen (Vukovic & Siegel, 2010). Somit kann gefolgert werden, dass es sich hier nicht zwingend um ein stabiles Phänomen handelt. Es gibt somit Kinder, die zumindest vorübergehend auf spezielle Unterstützung im Bereich Mathematik angewiesen sind, die aber nicht eine im klassischen Sinne andauernde Entwicklungsstörung haben (Mazzocco & Räsänen, 2013).

2.1.2 DSM-V

Definition von Rechenstörung

Im DSM-V (Diagnostic and Statistical Manual of Mental Disorders, deutsche Ausgabe von Falkai & Wittchen, 2015) wurden im Vergleich zum DSM-IV Überarbeitungen vorgenommen, die für das vorliegende Thema von Bedeutung sind. Die schulischen Entwicklungsstörungen, die in der ICD-10 nach Bereichen getrennt sind, werden in diesem Manual in der Kategorie spezifische Lernstörungen zusammengefasst (Falkai & Wittchen, 2015, S. 87–98; vgl. Schulte-Körne, 2014). Dabei müssen Schwierigkeiten in den Bereichen Lesen, Rechtschreiben oder Rechnen vorliegen (Schulte-Körne, 2014). Für die Diagnose einer Rechenstörung müssen sich spezifische Schwierigkeiten in den Teilbereichen Zahlenverständnis, arithmetisches Faktenwissen, genaues bzw. flüssiges Rechnen oder mathematisches Schlussfolgern zeigen (Schulte-Körne, 2014). Der Begriff der Dyskalkulie wird als alternative Bezeichnung aufgelistet, wobei darunter spezifisch Schwierigkeiten in der Verarbeitung von Zahlen, dem Erwerb arithmetischen Faktenwissens und dem genauen und schnellen Durchführen von Rechenoperationen verstanden werden (Schulte-Körne, 2014). Letzteres entspricht weitestgehend dem Verständnis der Rechenstörung gemäss ICD-10. Im DSM-V wird neu zwischen drei unterschiedlichen Schweregraden von Lernstörungen unterschieden:

> Ein leichter Schweregrad wird codiert, wenn einzelne Schwierigkeiten in einem oder zwei Lernbereichen aufgetreten sind, die aber noch kompensiert werden können. Dies setzt jedoch ausreichende Unterstützungsmaßnahmen voraus. Wenn trotz intensiver Förderung deutliche Schwierigkeiten in einem oder zwei Lernbereichen vorliegen, soll ein mittelgradiger Schweregrad codiert werden. Sind […] Anforderungen im Alltag nur sehr eingeschränkt zu bewältigen und können ohne intensive Hilfen und länger andauernde Unterstützungsmaßnahmen diese nicht gelöst werden, soll der Schweregrad schwer codiert werden. (Schulte-Körne, 2014, S. 370)

Weiter werden im DSM-V neu vier statt drei diagnostische Kriterien (A, B, C, D) ausführlich beschrieben (Schulte-Körne, 2014). So müssen sich entsprechende Schwierigkeiten trotz gezielter Interventionen über einen Zeitraum von mindestens sechs Monaten zeigen, damit eine Lernstörung diagnostiziert werden kann (Kriterium A; Falkai & Wittchen, 2015). Bedeutende Änderungen beziehen sich zudem auf die Erweiterung der einzusetzenden Instrumente für eine Diagnosestellung, mit der stark unterdurchschnittliche Leistungen festgestellt werden sollen (Kriterium B; Schulte-Körne, 2014). Hierfür wird zwar weiterhin empfohlen, standardisierte Testverfahren zu verwenden, zusätzlich können aber auch Vorgehen aus der klinischen Diagnostik hinzugezogen werden (z.B. Entwicklungs- und Familienanamnese, klinische Interviews, Schulberichte, Beurteilungsskalen sowie pädagogische oder psychologische Testdiagnostik; Schulte-Körne, 2014, S. 370). Einen weiteren, bedeutsamen Unterschied im Vergleich zur ICD-10 besteht in Bezug auf den Grenzwert für die Diagnosestellung bei Verwendung standardisierter Testverfahren. So kann gemäss DSM-V bereits ab einer Standardabweichung von 1.5 im Vergleich zur Normstichprobe in einem standardisierten Rechentest eine Diagnose vergeben werden. Beim Vorliegen mehrerer klinischer Hinweise auf eine Lernstörung wird zudem bereits eine Diskrepanz von einer Standardabweichung als ausreichend angesehen (Falkai &

Wittchen, 2015, S. 91). Ähnlich wie in der ICD-10 müssen auch gemäss DSM-V die weiter oben genannten Symptome bereits früh, d.h. in den ersten Schuljahren auftreten. Dabei wird jedoch stärker berücksichtigt, dass kompensatorische Faktoren dazu beitragen können, dass die Lernstörung erst später erkannt wird (Kriterium C; Schulte-Körne, 2014). Unter dem Kriterium D werden schliesslich noch Ausschlusskriterien genannt, die denen der ICD-10 ähnlich sind:

> Hierzu gehören eine Intelligenzbeeinträchtigung (IQ < 70), allgemeine Entwicklungsverzögerung und allgemeine Beeinträchtigung der Lernfähigkeit, Seh- und Hörstörungen, neurologische Erkrankungen und motorische Störungen. Außerdem sind Lernstörungen aufgrund fehlender Unterrichtung, von Schulabsentismus, in Folge von ökonomischen Faktoren der Familie oder Folge von nicht-förderlichen Umweltbedingungen ausgeschlossen. (Schulte-Körne, 2014, S. 371)

Der bedeutendste Unterschied zwischen ICD-10 und DSM-V ist der Verzicht auf das IQ-Diskrepanzkriterium (Schulte-Körne, 2014). D.h., die Leistung in einem standardisierten Test muss zwar ausserhalb der oben genannten Grenzwerte im Vergleich zu Schuljahrgangs- und Altersnormen liegen, es muss jedoch keine Diskrepanz zur allgemeinen Intelligenz mehr festgestellt werden, sodass die Erfassung der Intelligenz einzig für den Ausschluss einer allgemeinen Intelligenzminderung notwendig ist (Falkai & Wittchen, 2015). Somit wird in der überarbeiteten Version DSM-V die vielfältige Kritik am IQ-Diskrepanzkriterium aufgenommen und auf dieses Kriterium verzichtet, was angesichts der vielfältigen Kritikpunkte positiv gewertet werden kann. Es muss jedoch beachtet werden, dass in früheren Versionen des DSM das Vorliegen einer IQ-Diskrepanz als Kriterium verlangt wird, was für die Einschätzung von Studien, die auf älteren DSM-Definitionen beruhen, relevant ist (Saß, Wittchen, Zaudig & Houben, 2003).

2.1.3 Definitionen von Rechenschwäche

Wie im Folgenden aufgezeigt, wird der Begriff der Rechenschwäche im Gegensatz zum Begriff der Rechenstörung meistens für ein grösseres Spektrum von Schwierigkeiten beim Mathematiklernen verwendet, wobei jedoch teilweise unterschiedliche Definitionen zur Anwendung kommen. So wird der Begriff der Rechenschwäche in diversen Publikationen dann verwendet, wenn Kinder zwar im Sinne der ICD-10 gravierende mathematische Schwierigkeiten aufweisen, jedoch das zweite Diskrepanzkriterium (Mathematikleistung vs. Intelligenz) nicht erfüllt wird (Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie, 2007; Fischbach et al., 2013; Kuhn et al., 2013; Thomas et al., 2015). Entsprechend sollten bei diesem Begriffsverständnis gemäss obigen Ausführungen zur ICD-10-Definition bezüglich der Mathematikleistung ebenfalls strenge Cut-off-Werte vom 10. bis zum 16. Prozentrang angewendet werden. Es gibt jedoch auch Autorinnen und Autoren, die den Begriff der Rechenschwäche dann verwenden, wenn beim zweiten Diskrepanzkriterium ein weniger striktes Kriterium von einer Standardabweichung von 1.5 verlangt wird und somit nach wie vor eine zur Intelligenz diskrepante Mathematikleistung vorliegen muss (z.B. Jacobs & Petermann, 2005).

In vielen Studien wird mit dem Begriff der Rechenschwäche jedoch auf höhere Cut-off-Werte in Bezug auf die Mathematikleistung verwiesen, wobei eine Diskrepanz zur Intelligenz nicht als Diagnosekriterium verlangt wird (Mazzocco, 2005). So finden sich in Abgrenzung zum Begriff der Rechenstörung in Publikationen der Forschungsgruppe von Geary unter dem Begriff „low achievers in mathematics" Schülerinnen und Schüler mit Mathematikleistungen, die im untersten Quartil der Leistungen (25. Perzentil), jedoch über dem 10. Prozentrang liegen (z.B. Geary, Hoard, Nugent et al., 2012; Geary, Hoard, Byrd-Craven, Nugent & Numtee, 2007). Das 25. Perzentil wird auch in anderen Publikationen als Cut-off-Wert verwendet (z.B. Ennemoser & Krajewski, 2007; Jiménez Gonzalez & Garcia Espinel, 1999; McLean & Hitch, 1999). Teilweise werden aber auch weit höhere Cut-off-Werte bis zum 35. Perzentil eingesetzt (z.B. Jordan et al., 2003; Gersten, Jordan & Flojo, 2005). In all diesen Ansätzen stehen somit Kinder im Fokus, die ein breites Spektrum von mathematischen Schwierigkeiten aufweisen. Es handelt sich dabei um Kinder, deren Leistungen am unteren Ende eines als kontinuierlich anzusehenden Spektrums liegen und die (insbesondere bei grosszügigen Cut-off-Werten) allenfalls nur über zeitlich befristete Schwierigkeiten im Fach Mathematik verfügen (Mazzocco, 2005; Vukovic & Siegel, 2010). Andere Publikationen verlangen hingegen für das Vorliegen einer Rechenschwäche Mathematikleistungen, die sich bei mehreren Messzeitpunkten im untersten Quartil befinden, so dass sich die Definition von Rechenschwäche somit auf länger anhaltende mathematische Schwierigkeiten bezieht (z. B. Vanbist et al., 2014). Von anderen Autoren wird als ausschlaggebendes Kriterium für eine Rechenschwäche der Besuch von speziellen Förderlektionen im Fach Mathematik verlangt mit der Begründung, dass dadurch längerfristige mathematische Schwierigkeiten erfasst werden können (Andersson, 2010). Rechenschwäche wird in gewissen Publikationen aber auch allgemein „[…] als Versagen im Mathematikunterricht" bei unterdurchschnittlicher Mathematikleistung definiert, wobei in der genannten Publikation eine Leistungsabweichung von minus einer Standardabweichung verlangt wird (Moser Opitz, 2013, S. 139). Es kann somit festgehalten werden, dass auch beim Begriff der Rechenschwäche keine einheitliches Begriffsverständnis vorliegt, damit aber meist ein grösseres Spektrum an mathematischen Schwierigkeiten im Vergleich zu den Definitionen einer Rechenstörung gemäss ICD-10 oder DSM-V gemeint ist.

Bei der Verwendung des Begriffs Rechenschwäche wird zudem meist eine multikausale Verursachung der mathematischen Schwierigkeiten angenommen, wobei auch soziale, emotionale und unterrichtliche Faktoren stärker berücksichtigt werden und davon ausgegangen wird, dass je nach Kind verschiedene Faktoren an der Entstehung einer Rechenschwäche beteiligt sein können (z.B. Moser Opitz, 2013). Bei der Verwendung des Begriffs Rechenstörung finden sich hingegen eher Positionen, in denen versucht wird die mathematischen Schwierigkeiten primär durch Defizite in kognitiven Funktionen zu erklären, wobei diese sowohl allgemeine kognitive Funktionen (z.B. Arbeitsgedächtnis) als auch spezifisch mathematische kognitive Funktionen (z.B. Zahlensinn) umfassen können (z.B. Landerl & Kaufmann, 2008). Eine ausführliche Darstellung dieser Faktoren findet sich in Kapitel 2.3.

2.1.4 Prävalenz

Bei Prävalenzschätzungen muss beachtet werden, dass diese immer abhängig davon sind, welche Definition und somit welche Grenzwerte für die Feststellung einer Rechenstörung oder Rechenschwäche gewählt werden (Hasselhorn & Schuchardt, 2006; Wyschkon, Kohn, Ballaschk & Esser, 2009). Zudem können z.B. auch das Alter der untersuchten Kinder, die verwendeten Leistungstests oder die (fehlende) Berücksichtigung von Lese-/Rechtschreibleistungen einen Einfluss auf die festgestellten Prävalenzraten haben (von Aster et al., 2007; Wyschkon et al., 2009).

Wie Wyschkon et al. (2009) eindrücklich zeigen konnten, führt eine strenge Befolgung der ICD-10-Definition von Rechenstörung zu einer sehr tiefen Prävalenzrate von 0.5% (vgl. Tabelle 1). Fischbach et al. (2013) haben in einer deutschen Stichprobe am Ende der zweiten und am Anfang der dritten Klasse ebenfalls untersucht, wie hoch die Prävalenz von Rechenstörung unter Berücksichtigung des doppelten Diskrepanzkriteriums ist. Dabei wurde bereits eine Diskrepanz von einer Standardabweichung und somit ein Prozentrang kleiner 16 als ausreichend definiert. Unter Berücksichtigung dieses Grenzwertes wiesen 2.6% der untersuchten Kinder eine isolierte Rechenstörung und 2% eine kombinierte Lernstörung (Rechenstörung und Lese-/Rechtschreibstörung) auf. Im Kontrast dazu lagen die Prävalenzraten höher, wenn das zweite Diskrepanzkriterium nicht berücksichtigt wurde. Eine isolierte Rechenschwäche wurde so bei 5.0% der untersuchten Kinder festgestellt, und 4.2% der Kinder wiesen eine kombinierte Lernschwäche auf (Fischbach et al., 2013). Weiter wurde auch berechnet, wie hoch die Prävalenz liegt, wenn bei der isolierten Rechenstörung und der isolierten Rechenschwäche auf eine Kontrolle der Lese-/Rechtschreibfähigkeiten verzichtet wird. Auch dies führt zu höheren Raten von 5.0% für eine Rechenstörung und 9.2% für eine Rechenschwäche (Fischbach et al., 2013).

Tabelle 1: Prävalenzraten für Rechenstörungen und Rechenschwäche sowie für kombinierte Lernstörungen/-schwächen (in Prozent) gemäss ausgewählten Studien[3]

Studie	Isolierte Rechenstörung	Kombinierte Lernstörung	Isolierte Rechenschwäche	Kombinierte Rechenschwäche
Wyschkon et al., 2009 (ICD-10-Kriterien)	0.5	–	–	–
Fischbach et al., 2013	2.6	2.0	5.0	4.2
Wyschkon et al., 2009, basierend auf Kriterien von Jacobs & Petermann, 2003	3.9	–	–	–
Wyschkon et al., 2009, basierend auf Kriterien von Esser, Wyschkon & Ballaschk, 2008	2.5	–	–	–
Wyschkon et al., 2009, basierend auf Kriterien von von Aster et al., 2007	1.6	–	–	–
Wyschkon et al., 2009, basierend auf Kriterien von Gross-Tsur, Manor & Shalev, 1996	3.4	–	–	–
Wyschkon et al., 2009, basierend auf Kriterien von Lewis, Hitch & Walker, 1994	4.5	–	–	–
Wyschkon et al., 2009, basierend auf LRS-Kriterien von Schulte-Körne, Deimel & Remschmidt, 2001	4.8	–	–	–
Wyschkon et al., 2009, basierend auf Kriterien von Fuchs et al., 2005	8.1	–	–	–
Lewis et al., 1994	1.3	2.3	–	–
von Aster et al., 2007	1.8	4.2	–	–

In der Untersuchung von Wyschkon et al. (2009) wurden weitere Definitionskriterien auf die bereits oben erwähnte Stichprobe angewendet. Hier zeigte sich für zwei Diagnosekriterien, die das doppelte Diskrepanzkriterium berücksichtigt haben, aber etwas weniger strenge Grenzwerte als bei der ICD-10 verwendeten, dass 3.9% (basierend auf den Kriterien von Jacobs & Petermann, 2003) sowie 2.5% (basierend auf den Kriterien von Esser et al., 2008) der untersuchten Kinder eine isolierte Rechenstörung haben (Wyschkon et al., 2009). Weiter wurden von Wyschkon et al. (2009) auch Kriterien für die Diagnosestellung verwendet, bei denen ein einfaches Diskrepanzmodell, eine Minderleistung bei normalem IQ oder ein Regressionsmodell zur Anwendung kamen, das an dieser Stelle nicht weiter ausgeführt werden soll. Dabei ergaben sich sowohl eine sehr niedrige Prävalenzrate von 1.6% (basierend auf Grenzwerten von von Aster et al., 2007), eine mittlere Prävalenzraten von 3.4 bis 4.8% (basierend auf den Grenzwer-

[3] Es wird in der Tabelle jeweils derjenige Begriff verwendet (Rechenstörung vs. Rechenschwäche), der auch in der jeweiligen Studie zu finden war.

ten von Gross-Tsur et al., 1996, Lewis et al., 1994 und dem für die Diagnose einer Lese-/ Rechtschreibstörung entwickelten Regressionsmodells von Schulte-Körne et al., 2001) als auch eine als eher hoch einzuschätzende Prävalenzrate von 8.1% (basierend auf den Grenzwerten von Fuchs et al., 2005).

In der Studie von Lewis et al. (1994) wurde im Vergleich zur Stichprobe von Wyschkon et al. (2009) unter Verwendung der gleichen Diagnosekriterien eine deutlich niedrigere Prävalenz von 1.3% festgestellt. Zusätzlich wurde für die kombinierte Lernstörung eine Prävalenz von 2.3% berechnet (Lewis et al., 1994). Ein möglicher Grund für den Unterschied zwischen den beiden Studien könnte in den verwendeten Testverfahren für die Mathematikleistung liegen. Lewis et al. (1994) haben einen Test verwendet, der einzig Rechenfertigkeiten erfasst, wohingegen beim Test von Wyschkon et al. (2009) zusätzlich auch Sachaufgaben, Aufgaben zum Mengenvergleich und zum Zahlenstrahl, das Erkennen von Formen und Uhrzeit sowie Rundungsaufgaben eingesetzt wurden. Es kann daher vermutet werden, dass Testverfahren, die stärker auf Rechenfertigkeiten fokussieren, dazu führen, dass bei weniger Kindern eine Rechenstörung diagnostiziert wird. Kaum Abweichungen gab es hingegen zwischen den Berechnungen von Wyschkon et al. (2009) und von von Aster et al. (2007). Letztere haben in ihrer Studie Schülerinnen und Schüler des zweiten Schuljahres in der Deutschschweiz untersucht, dabei auf das Intelligenz-Diskrepanz-Kriterium verzichtet und für die Mathematikleistung eine Diskrepanz von 1.5 Standardabweichungen zum Mittelwert des standardisierten Mathematiktests (ZAREKI-R; von Aster, Weinhold Zulauf & Horn, 2006) verwendet. Zudem durfte für eine isolierte Rechenstörung die Lese- und Rechtschreibleistung nicht mehr als 0.5 Standardabweichungen vom Mittelwert der verwendeten Tests abfallen. Dabei ergaben sich Prävalenzraten von 1.8% für die isolierte Rechenstörung und 4.2% für die kombinierte Lernstörung. Letzterer Wert ist im Vergleich zu Fischbach et al. (2013) und Lewis et al. (1994) als eher hoch einzuschätzen.

Wie die vorher genannten Studien aufzeigen, können keine allgemeinen Aussagen zur Prävalenz von Rechenstörung bzw. Rechenschwäche gemacht werden. Je nach verwendeten Kriterien für die Diagnose einer Rechenstörung bzw. einer Rechenschwäche können diese sehr unterschiedlich ausfallen. Ebenso muss berücksichtigt werden, dass auch die verwendeten Testverfahren eine Auswirkung auf die Höhe der Prävalenzschätzung haben können. Grundsätzlich kann jedoch gesagt werden, dass bei der Anwendung strenger Grenzwertkriterien niedrigere Prävalenzraten resultieren, als wenn höhere Grenzwerte zum Einsatz kommen. Auch bei den Prävalenzschätzungen zeigt sich somit die Problematik, dass im wissenschaftlichen Diskurs bisher keine Verständigung auf eine einheitliche Definition zustande gekommen ist (Murphy et al., 2007).

In der vorliegenden Arbeit wird im Folgenden der Begriff der Rechenschwäche für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen verwendet. Dieser Begriff findet sich inzwischen in diversen neueren Publikationen aus dem deutschsprachigen Raum (vgl. z.B. Busch, Oranu, Schmidt & Grube, 2013; Ennemoser et al., 2015; U. Fischer, Roesch & Moeller, 2017; Häsel-Weide, 2016; Kohn et al., 2013; Moser Opitz & Ramseier, 2012; Rauscher et al., 2017; Wißmann et al., 2013). Wie im Kapitel 2.1 dargestellt, sind die gängigen Cut-off-Werte, die bei diesem Begriff jeweils zum Einsatz kommen, teilweise sehr unterschiedlich. Wenn es in der vorliegenden Arbeit um die Analyse der Ergebnisse der zu untersuchenden Interventionsstudie geht, wird für die Abgrenzung von unterdurchschnittlichen im Vergleich zu durchschnittlichen Mathematikleistungen als Cut-off-Wert das 25. Perzentil gewählt. Auf den Begriff der Rechenstörung wird in dieser Arbeit hingegen nur noch dann zurückgegriffen, wenn es explizit um eine Abgrenzung zwischen Kindern geht, die das doppelte Diskrepanzkriterium einer Rechenstörung erfüllen, und Kindern, die unabhängig davon unterdurchschnittliche Mathematikleistungen aufweisen.

2.2 Modelle mathematischer Verarbeitung und Entwicklung

Zwar gibt es bis heute nur wenige umfassende Modelle mathematischer Entwicklung (vgl. W. Schneider, Küspert & Krajewski, 2013), diese beinhalten jedoch wichtige Informationen zu zentralen Entwicklungsschritten beim Aufbau des Zahlbegriffs und zu mathematische Inhalten und Kompetenzen, die für die weitere mathematische Entwicklung bedeutsam sind. Modelle der mathematischen Verarbeitung wiederum gehen der Frage nach, wie Zahlen kognitiv verarbeitet und mental repräsentiert werden, und stellen somit eine wichtige Basis zum Verständnis mathematischer Kompetenzen und deren Entwicklung dar. In diesem Kapitel werden daher Modelle der mathematischen Entwicklung und Verarbeitung vorgestellt, die für die inhaltliche Gestaltung der in dieser Arbeit zu untersuchenden unterrichtsintegrierten Mathematikförderung hilfreiche Hinweise bieten. Nicht thematisiert werden im Folgenden Modelle, die sich nur mit einem einzelnen Aspekt der mathematischen Entwicklung auseinandersetzen wie z.B. das Modell der Zählentwicklung von Fuson (1988).

2.2.1 Das Triple-Code-Modell von Dehaene und Cohen

Die Verarbeitung von Zahlen gilt grundsätzlich als eine hochkomplexe kognitive Leistung, die bis heute erst teilweise verstanden wird und aus verschiedenen Teilkomponenten zusammengesetzt ist (W. Schneider et al., 2013). Dennoch gibt es einige Modelle, die wichtige Erkenntnisse zur Zahlenverarbeitung zusammenfassen. Als eines der zentralsten Modelle wird das Triple-Code-Modell von Dehaene (1992) und Dehaene und Cohen (1995) angesehen. Weitere wichtige Modelle zur Verarbeitung von Zahlen stammen von McCloskey, Caramazza und Basili (1985; Modell der Zahlenverarbeitung und des Rechnens), Noël und Seron (1993; Modell des bevorzugten Eingangsmodus) und Cipolotti und Butterworth (1995; Multiples Transkodiermodell), diese sollen hier jedoch nicht weiter ausgeführt werden.

Das Triple-Code-Modell beschreibt, wie Zahlen mental verarbeitet werden und welche Hirnregionen bei der Zahlenverarbeitung involviert sind. Es wird angenommen, dass es drei Module der Zahlenverarbeitung gibt, die visuell-arabische Zahlenform, die verbal-phonologische Zahlenform und die analoge Grössenrepräsentation (vgl. Abbildung 1).

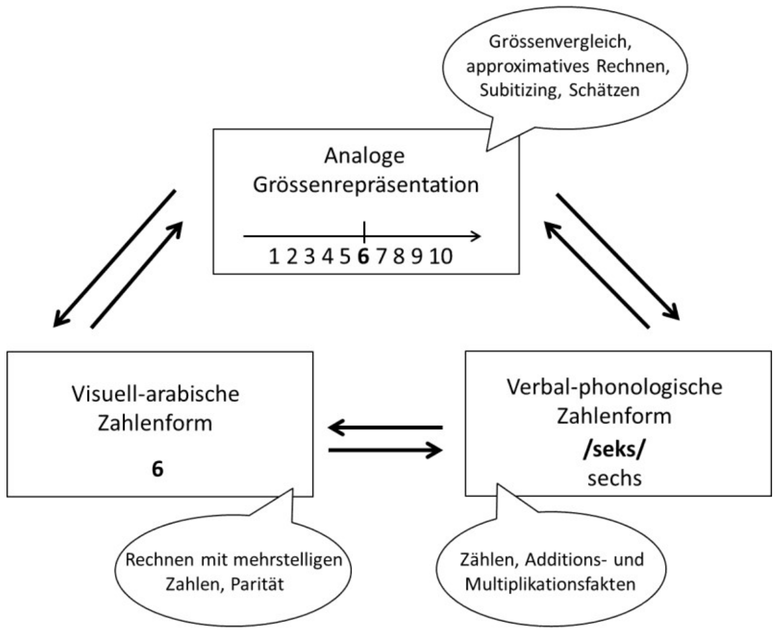


Abbildung 1: Triple-Code-Modell gemäss Dehaene (1992), leicht modifiziert nach Landerl und Kaufmann (2008, S. 25)

Gemäss diesem Modell werden Zahlen in unterschiedlicher Form im Gehirn repräsentiert: Ziffern im *Modul der visuell-arabischen Zahlenform* und Zahlwörter im *Modul der verbal-phonologischen Zahlenform* (Dehaene & Cohen, 1995). Die Grösse bzw. Mächtigkeit einer Zahl (Zahlensemantik) hingegen erschliesst sich im *Modul der analogen Grössenrepräsentation*. Im Triple-Code-Modell wird davon ausgegangen, dass jederzeit Zahlen von einer Repräsentationsform in die andere transkodiert werden können, was insbesondere bei komplexeren mathematischen Aufgabenstellungen unerlässlich ist (Dehaene & Cohen, 1995). Grössenvergleiche von Zahlen oder das Schätzen von Mengen werden in diesem Modell dem Modul der analogen Grössenrepräsentation zugeteilt, Zählen und arithmetisches Faktenwissen hingegen werden als verbal-phonologische Zahlenform verarbeitet, beim schriftlichen Rechnen mit grossen Zahlen steht schliesslich die visuell-arabische Zahlenform im Vordergrund (Dehaene & Cohen, 1995).

Die Lokalisation der einzelnen Module im Gehirn basiert auf Untersuchungen an erwachsenen Menschen mit erworbenen Rechenstörungen mit spezifischen Einschränkungen in den mathematischen Kompetenzen, die auf Beschädigungen in bestimmten Hirnarealen zurückgeführt werden konnten (Dehaene & Cohen, 1995). Das Modul der analogen Grössenrepräsentation und das Modul der visuell-arabischen Zahlenform sind sowohl links- als auch rechtshemisphärisch zu finden, das verbal-phonologische Zahlenmodul hingegen wird nur der linken Hirnhemisphäre zugeordnet, die auch für die Verarbeitung von Sprache verantwortlich ist (Dehaene & Cohen, 1995). Genauer befindet sich das Modul der analogen Grössenrepräsentation im sogenannten intraparietalen Sulcus (ISP), welcher Teil des Parietallappens ist, die verbal-phonologische Zahlenform lässt sich anatomisch der perisylvischen Furche und Basalganglien in subkortikalen Hirnregionen zuordnen, und die visuell-arabische Zahlenform wird mit okzipitalen Hirnregionen in Verbindung gebracht (Dehaene & Cohen, 1995).[4] Zusätzlich zu den genannten parietalen Hirnregionen sind auch frontale Hirnregionen insbesondere bei komplexen Rechenleistungen von grosser Bedeutung (Dehaene & Cohen, 1995). Dort ist das Arbeitsgedächtnis lokalisiert, das in Zusammenhang mit Rechenschwäche vielfach diskutiert und im Kapitel 2.3.1 noch ausführlich thematisiert wird. Der Mehrwert des Triple-Code-Modells liegt darin, dass es die Verarbeitung der drei Repräsentationsformen von Zahlen konzeptuell beschreiben kann und die Bedeutsamkeit der Fähigkeit zwischen den verschiedenen Repräsentationsformen flexibel hin und her wechseln zu können betont. Zu beachten ist jedoch, dass sich dieses Modell in erster Linie auf die Verarbeitung von Zahlen bei erwachsenen Menschen bezieht und bei Kindern daher gewisse Abweichungen nicht auszuschliessen sind (Landerl & Kaufmann, 2008).

2.2.2 Das Vier-Stufen-Entwicklungsmodell der Zahlenverarbeitung von von Aster et al.

Beim Vier-Stufen-Entwicklungsmodell der Zahlenverarbeitung (von Aster, Kucian, Schweiter & Martin, 2005; von Aster, 2005) handelt es sich um ein Entwicklungsmodell mit stark neuropsychologischem Fokus, das auf dem Triple-Code-Modell von Dehaene (1992) und Dehaene und Cohen (1995) aufbaut.

[4] Für die deutsche Übersetzung der englischen Fachbegriffe in Bezug auf die hirnanatomische Lokalisation der Module wurde auf Landerl und Kaufmann (2008) zurückgegriffen.

Im Vier-Stufen-Entwicklungsmodell werden auf der *ersten Stufe* Fähigkeiten verordnet, die als angeboren angesehen werden (vgl. Abbildung 2). Solche angeborenen numerischen Grundkompetenzen werden als „core system“ bezeichnet und als angeborener Zahlen- oder Mengensinn verstanden (Feigenson, Dehaene & Spelke, 2004; von Aster, Kucian & Martin, 2006). Demnach können bereits Babys bei kleinen Mengen deren kardinale Grösse erfassen („subitizing“) und zwei Mengen aufgrund ihrer Grösse voneinander unterscheiden, wenn der Grössenunterschied ausreichend stark ausgeprägt ist (von Aster, Kucian et al., 2006). Die *zweite Stufe* bezieht sich auf das Kleinkind- und Vorschulalter (von Aster et al., 2005). Gemäss dem Vier-Stufen-Modell eignen sich „[…] Kinder mit Beginn der Sprachentwicklung die Zahlwortsequenz, Zählprinzipien (z.B. Eins-zu-Eins-Zuordnung, stabile Reihenfolge), das Zu- und Wegzählen zum Verändern von Mengen und Begriffe wie ‚mehr‘ und ‚weniger‘ […]“ an (von Aster et al., 2005, S. 614). Zudem können mithilfe von Zählstrategien bereits erste, einfache Additions- und Subtraktionsaufgaben gelöst werden.

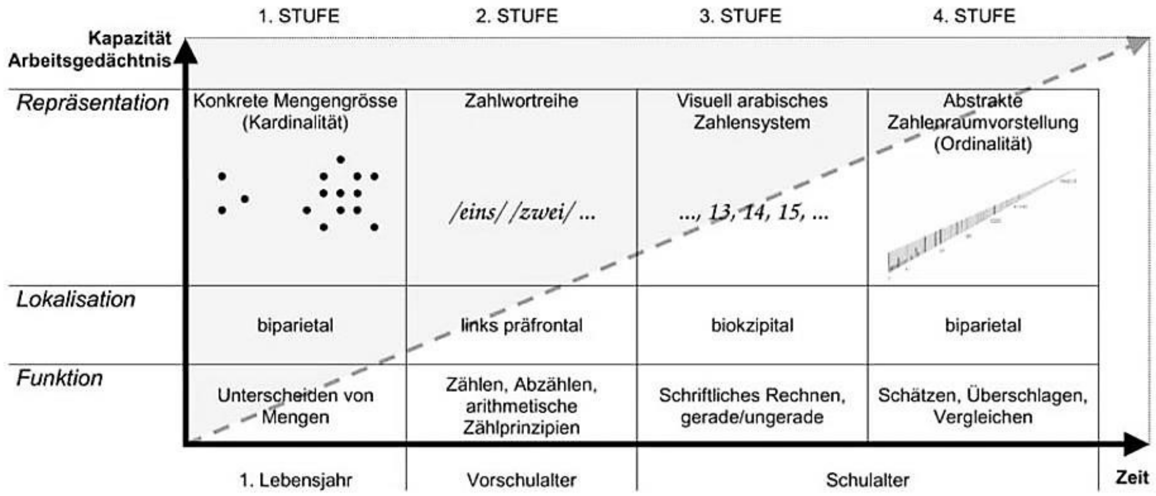


Abbildung 2: Vier-Stufen-Entwicklungsmodell der Zahlenverarbeitung (von Aster, Kucian et al., 2006, S. 3)

Auf der *dritten Stufe* des Entwicklungsmodells eignen sich Kinder das arabische Notationssystem für Zahlen an und müssen dabei auch „kulturspezifische Übersetzungsregeln“ (von Aster et al., 2005, S. 614) erlernen (z.B. Zehner-Einer-Inversion im deutschsprachigen Raum). Auf der *vierten Stufe* entwickelt sich eine innere abstrakte Zahlenraum- bzw. Zahlenstrahlvorstellung, Diese mentale Zahlenraumvorstellung ermöglicht es, Zahlen nach ihrer Grösse zu vergleichen, „arithmetisch zu manövrieren“, Aufgaben zu schätzen und zu runden (von Aster et al., 2005, S. 616).

Das Vier-Stufen-Entwicklungsmodell legt dar, wie Kinder eine mentale Zahlenraumvorstellung aufbauen, indem sie sich ausgehend von einer konkreten Mengenrepräsentation die verbale Zahlwortreihe und daran anschliessend das visuell-arabische Zahlensystem aneignen, was dann in eine abstrakte Zahlenraumvorstellung mündet (W. Schneider et al., 2013). Das Modell stellt somit den Versuch dar, die drei Zahlenrepräsentationsformen des Triple-Code-Modells von Dehaene (1992) und Dehaene und Cohen (1995) in ein Entwicklungsmodell zu übertragen. Allerdings bleibt das Modell recht allgemein, wenn es um die Beschreibung der Funktionen geht, die mit den jeweiligen Repräsentationsstufen in Verbindung gebracht werden (vgl. Abbildung 2). Dadurch kann das Modell nur in Ansätzen beschreiben, wie die Entwicklung der einzelnen Repräsentationsformen im Detail mit der Entwicklung der konkreten mathematischen Kompetenzen in Verbindung steht, stellt aber dennoch eine wichtige Ergänzung zum Triple-Code-Modell dar, da es eine Entwicklungsperspektive einnimmt.

2.2.3 Fünfstufiges Modell der mathematischen Entwicklung von Fritz und Ricken

Das Modell von Fritz und Ricken (2008) basiert auf einer entwicklungspsychologischen Perspektive. Es baut u.a. auf theoretischen und empirischen Erkenntnissen von Fuson (1988), Resnick (1983), Wynn (1990), Piaget und Szeminska (1975) und Case und Okamoto (1996) auf und bezieht sich auf den Altersbereich von vier bis acht Jahren (Langhorst, Hildenbrand, Ehlert, Ricken & Fritz, 2013).

Auf der *ersten Stufe* der mathematischen Entwicklung (Erwerb der Zählzahl*)* eignen sich Kinder die Zahlwortreihe an, wobei zu Beginn noch kein eigentliches Zählen, sondern ein auswendig gelerntes Wiedergeben im Sinne eines zusammenhängenden „Wortgebildes" stattfindet (Fritz & Ricken, 2008; Ricken, Fritz & Balzer, 2011). Zudem können auf dieser Stufe durch Eins-zu-eins-Zuordnung erste Mengenvergleiche vorgenommen und mit der Zeit auch kleine Mengen ausgezählt werden (Fritz & Ricken, 2008; Langhorst et al., 2013). Auf der *zweiten Stufe,* der Repräsentation des mentalen Zahlenstrahls, wird die Zahlwortreihe nach und nach als aufsteigende Sequenz verstanden, und zwei Zahlen können mithilfe ihrer Position auf dem mentalen Zahlenstrahl miteinander verglichen werden (Ricken et al., 2011). Zudem können auf dieser Stufe Zählakte durchgeführt werden, diese müssen aber jeweils beim Zahlwort eins beginnen, da die Zahlwörter noch als feste Sequenz erfasst werden (Fritz & Ricken, 2008). Ebenfalls können erste Additionen und Subtraktionen zählend gelöst werden, indem auf dem mentalen Zahlenstrahl vor- oder zurückgegangen wird, allerdings wird das Ergebnis einer solchen Rechnung noch nicht als Summe, sondern als eine bestimmte Position auf dem Zahlenstrahl verstanden (Langhorst et al., 2013). Erst auf der *dritten Stuf*e, der Kardinalität und Zerlegbarkeit, wird erkannt, dass ein bestimmtes Zahlwort für eine bestimmte Anzahl von Objekten steht, was dazu führt, dass Zahlen aufgrund ihrer Mächtigkeit miteinander verglichen werden können (Langhorst et al., 2013). Zudem muss beim Zählen nicht mehr bei eins gestartet werden, sodass bei einer Aufgabe wie 4 + 3 = ? bei der Zahl vier startend direkt weitergezählt werden kann (Fritz & Ricken, 2008).

Die *vierte Stufe* des Enthaltenseins bezieht sich darauf, dass sich Kinder auf dieser Stufe ein differenzierteres Teile-Ganzes-Konzept aneignen (Langhorst et al., 2013). Sie wissen, dass eine Menge in verschiedene Teile zerlegt und wieder zu einem Ganzen zusammengefügt werden und dass aus dem Ganzen und einer Teilmenge die zweite Teilmenge bestimmt werden kann (Fritz & Ricken, 2008; Langhorst et al., 2013). Nach der Darstellung im Buch von Fritz und Ricken (2008) entwickelt sich auf der *vierten Stufe* auch bereits ein erstes Verständnis für den relationalen Zahlbegriff, d.h., es wird erkannt, dass der Abstand zwischen zwei aufeinanderfolgenden Zahlen immer eins beträgt. Auf der *fünften Stufe* werden gemäss dieser Quelle dann das Verständnis für das Teile-Ganzes-Konzept und der relationale Zahlbegriff vertieft und ausdifferenziert, was u.a. dazu führt, dass die Komplementarität der Addition und der Subtraktion erkannt wird (Fritz & Ricken, 2008). Gemäss der Darstellung des Modells in den Texten von Langhorst et al. (2013) und Ricken et al. (2011) wird die Entwicklung des relationalen Zahlbegriffs hingegen auf die *fünfte Stufe* verordnet, die entsprechend auch Relationalität genannt wird. Auf der fünften Stufe sind demnach dann auch Differenzbildungen zwischen zwei Zahlen möglich (Ricken et al., 2011).

Fritz und Ricken (2008) weisen darauf hin, dass die mathematische Entwicklung gemäss ihrem Entwicklungsmodell ein langfristiger, nicht linearer Prozess ist. Demnach können gewisse Einsichten bereits für kleinere Zahlenräume oder unter Verwendung von Material vorhanden sein, aber allenfalls noch nicht auf komplexere Sachverhalte übertragen werden. Das Modell überzeugt dadurch, dass verschiedene zentrale Erkenntnisse des bisherigen Theoriestandes zur mathematischen Entwicklung miteinander verknüpft wurden. Die abweichende Darstellung des Modells in verschiedenen Quellen, insbesondere die teilweise unterschiedliche Einordnung der Erkenntnisprozesse auf verschiedene Stufen, führt jedoch dazu, dass das Modell nicht immer eindeutig ist. Das Modell konnte im Rahmen der Validierung des MARKO-D-Tests (Ricken, Fritz-Stratmann & Balzer, 2013), der auf diesem Modell aufbaut, jedoch bestätigt werden. So zeigte sich in einer Studie, dass 90% aller mit dem MARKO-D untersuchten Kinder eindeutig einer der fünf Entwicklungsstufen des Modells von Fritz und Ricken (2008) zugeordnet werden konnten (Ricken et al., 2011). Kinder, die in höhere Entwicklungsstufen eingeteilt wurden, haben somit mehr Aufgaben korrekt lösen können und waren in Aufgaben, die niedrigere Entwicklungsniveaus abbilden, erfolgreicher als Kinder, die sich gemäss Testergebnis noch nicht auf der gleichen Entwicklungsstufe befanden (Ricken et al., 2011).

2.2.4 Entwicklungsmodell der Zahl-Grössen-Verknüpfung von Krajewski

Beim Modell der Zahl-Grössen-Verknüpfung von Krajewski (2003, 2005, 2007) handelt es sich ebenfalls um ein Modell mit entwicklungslogischer Perspektive, wobei auch hier die frühe mathematische Entwicklung im Zentrum steht. Als zentrale Bezugstheorie wird die Arbeit von Resnick (1989) zu „protoquantitativen Schemata“ genannt, weiter wurden auch Erkenntnisse von Piaget und Szeminska (1975), Gelman und Gallistel (1978) und Fuson (1988) einbezogen (W. Schneider et al., 2013). Im Fokus des Modells stehen Überlegungen zur Verknüpfung von Zahlen mit Grössenrelationen, die auf drei Ebenen verortet werden (Krajewski & Ennemoser, 2013; vgl. Abbildung 3). Das Modell wurde über die Jahre hinweg leicht umstrukturiert und gewisse Begrifflichkeiten angepasst (Krajewski & Ennemoser, 2013), daher wird an dieser Stelle auf aktuelle Quellen zurückgegriffen.

Auf *Ebene 1* werden erste Basisfertigkeiten entwickelt, die sich auf die Wahrnehmung von Mengen- bzw. Grössenunterschieden und die Aneignung von Zahlwörtern beziehen (W. Schneider et al., 2013). Ausgangspunkt für die Verknüpfung von Zahlen und Mengen ist die schon bei Säuglingen beobachtbare Fähigkeit, Mengen aufgrund ihrer Fläche, Ausdehnung und ihres Volumens zu unterscheiden (Feigenson, Carey & Spelke, 2002; Krajewski & Ennemoser, 2013). Dabei wird jedoch im Gegensatz zum Modell von von Aster et al. (2005) nicht angenommen, dass ein Verständnis für die Kardinalität einer Menge angeboren ist, sondern, dass es sich hierbei um eine rein visuelle Diskriminationsfähigkeit handelt (W. Schneider et al., 2013). Zudem eignen sich Kinder etwa ab dem zweiten Lebensjahr erste Zahlwörter an und lernen, diese als Zahlwortreihe aufzusagen, ohne dass sie jedoch die Zahlwörter bereits in Verbindung mit einer Mengenvorstellung bringen (Krajewski & Ennemoser, 2013).

Diese Verknüpfung geschieht erst auf *Ebene 2* in zwei verschiedenen Phasen. In der *Phase 2a* wird zunächst eine unpräzise Grössenrepräsentation bzw. ein unpräzises Anzahlkonzept aufgebaut, indem Kinder durch Alltagserfahrungen lernen, dass gewisse Zahlen mit „wenig“, andere Zahlen mit „viel“ assoziiert sind (W. Schneider et al., 2013). Zudem erleben sie beim Zählen, dass man für gewisse Zahlen nur „wenig“, für andere Zahlen „viel“ zählen muss, bis man bei einer bestimmten Zahl angelangt ist. Darauf basierend können erste Grössenunterscheidungen zwischen zwei Zahlen durchgeführt werden, wenn diese genügend weit auseinanderliegen (Krajewski & Ennemoser, 2013).

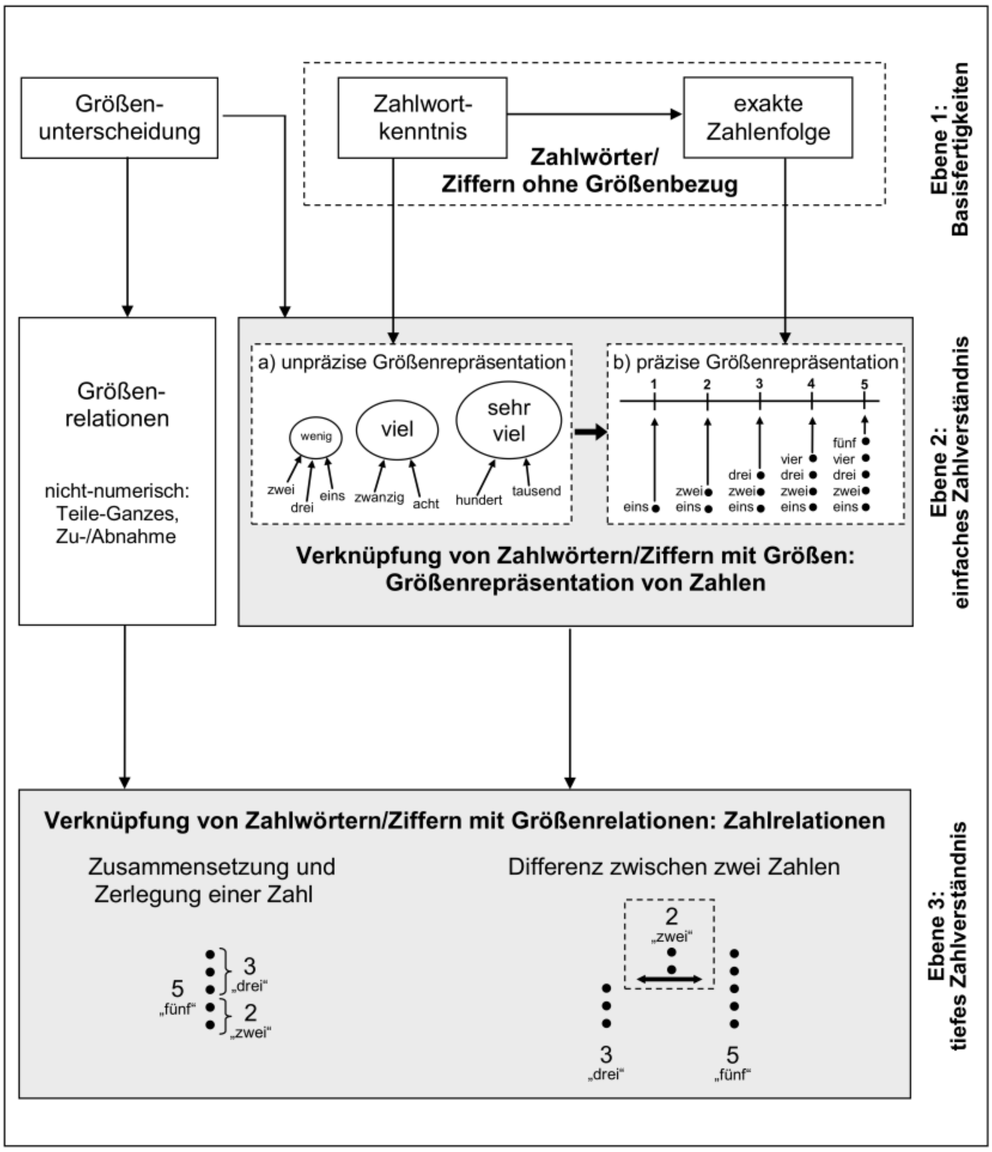


Abbildung 3: Entwicklungsmodell der Zahl-Grössen-Verknüpfung nach Krajewski (Krajewski & Ennemoser, 2013, S. 43)

In der *Phase 2b* des präzisen Anzahlkonzepts bzw. der präzisen Grössenrepräsentation erfolgen Lernprozesse, in denen das Konzept der Kardinalität aufgebaut wird, also ein Verständnis dafür, dass eine Zahl einer genauen Anzahl Dinge entspricht (Krajewski & Ennemoser, 2013). Auch hier sind Zählprozesse wichtig, da mithilfe dieser auch kleine Unterschiede bezüglich der Anzahl Zählschritte zwischen nahe beieinanderliegender Zahlen immer differenzierter erfasst werden können und so eine Verknüpfung zwischen Zahl und Anzahl aufgebaut werden kann (W. Schneider et al., 2013). Ein präzises Anzahlkonzept entsteht zuerst für kleinere Zahlenräume, dafür ist eine sichere Beherrschung der Zahlwortreihe (Ebene 1) im entsprechenden Zahlenraum notwendig (W. Schneider et al., 2013). Unabhängig von der Entwicklung im Bereich der Grössenrepräsentationen eignen sich Kinder auf *Ebene 2* auch wesentliche Konzepte im Bereich von Mengen- bzw. Grössenrelationen ohne Zahlbezug an (W. Schneider et al., 2013). „Im Zuge dieser Entwicklung lernen Kinder, dass eine diskrete Menge in kleinere Mengen zerlegt und aus diesen wieder zusammengesetzt werden kann (Teile-Ganzes) und begreifen, dass sich eine Menge nur dann verändert, wenn ihr etwas hinzugefügt oder weggenommen wird (Zu-/ Abnahme)“ (Krajewski & Ennemoser, 2013, S. 45).

Auf *Ebene 3* findet dann eine Verknüpfung der Zahlwörter bzw. der Zahlen mit den Mengen- bzw. Grössenrelationen statt. Kinder lernen zunächst für kleinere Zahlenräume, später auch für grössere, dass Zahlen aus kleineren Zahlen zusammengesetzt sind und jede Zahl in kleinere Zahlen zerlegt werden kann (Krajewski & Ennemoser, 2013). Zudem entwickelt sich auf dieser Ebene das Verständnis, dass die Differenz zwischen zwei Zahlen ebenfalls durch eine Zahl ausgedrückt werden kann (Krajewski & Ennemoser, 2013).

Krajewski und Ennemoser (2013, S. 45) halten fest, „[…] dass in Abhängigkeit von der mentalen wie auch externen Repräsentationsform sowie vom zugrunde liegenden Zahlenraum auch mehrere Entwicklungen parallel ablaufen und dabei deutlich gegeneinander verschoben sein können“. So können sich Kinder auf Ebene der Zahlwörter bereits auf einer höheren Entwicklungsebene befinden als unter Verwendung von Zahlen, oder sie können unter Zuhilfenahme von Arbeitsmaterialien oder in einem kleinen Zahlenraum bereits Aufgaben lösen, die sie ohne diese Hilfen oder in grösseren Zahlenräumen noch nicht bewältigen können (W. Schneider et al., 2013). Im Vergleich zum Modell von Fritz und Ricken (2008) ist das Modell von Krajewski (2003, 2005, 2008a) mit seinem Fokus auf die Verknüpfung von Zahlen und Grössenrepräsentationen als eindeutiger und klarer einzuschätzen, insbesondere auch mit Bezug auf die Abgrenzung zwischen den einzelnen Ebenen und deren Verknüpfungen. Allerdings wurde das Modell auch kritisiert, da einzelne Begriffsverwendungen aus mathematikdidaktischer Sicht als problematisch angeschaut werden (Nolte, 2015). Besonders die Verwendung des Begriffs „Grösse“ wird als missverständlich bewertet. Im mathematikdidaktischen Diskurs wird der Begriff der „Grösse“ als Verbindung von Masszahlen mit Einheiten verstanden: „Grössen sind z.B. Längen, Gewichte, Uhrzeiten etc.“ (Nolte, 2015, S. 63). Im Gegensatz dazu verwendet Krajewski (2003, 2005, 2008a) den Begriff der „Grösse“ für das kardinale Zahlenverständnis und meint somit die Verknüpfung von Zahlwörtern mit Mengenvorstellungen. Trotzdem wird das Modell inhaltlich als überzeugend bewertet (Nolte, 2015).

Die in diesem Kapitel thematisierten Modelle bieten einige Hinweise für die Gestaltung einer verstehensorientierten Mathematikförderung. Das Triple-Code-Modell (Dehaene, 1992; Dehaene & Cohen, 1995) zeigt auf, dass Zahlen unterschiedlich verarbeitet werden, dies gilt es im Unterricht zu berücksichtigen. Je nach Entwicklungsstand stehen zudem unterschiedliche Repräsentationsformen im Fokus (von Aster et al., 2005; von Aster, 2005). Ebenso muss berücksichtigt werden, dass sich Lernende auf unterschiedlichen Entwicklungsstufen befinden können, je nachdem in welchem Zahlenraum sie arbeiten bzw. um welche Repräsentationsform es sich handelt (Fritz & Ricken, 2008; Krajewski & Ennemoser, 2013). Zudem betonen die Modelle die Bedeutsamkeit von mathematischen Vorläuferfertigkeiten für die mathematische Entwicklung und weisen darauf hin, dass Schülerinnen und Schüler sich den kardinalen und den ordinalen Zahlaspekt ebenso aneignen müssen, wie ein Verständnis für das Teile-Ganzes-Konzept (Fritz & Ricken, 2008; Krajewski & Ennemoser, 2013; von Aster et al., 2005).

2.3 Einflussfaktoren

Nachdem im vorangehenden Kapitel zentrale Modelle der mathematischen Entwicklung und der numerischen Verarbeitung vorgestellt worden sind, wird im folgenden Kapitel der Frage nachgegangen, welche Faktoren bei der Entstehung einer Rechenschwäche eine Rolle spielen und wie diese die Mathematikleistung beeinflussen. Diese Faktoren bieten wichtige Hinweise für die methodische Gestaltung der in dieser Arbeit zu untersuchenden unterrichtsintegrierten Mathematikförderung und weisen auf zentrale Einflussfaktoren für die mathematische Leistung hin, die daher im Rahmen eines multikausalen Verständnisses von Mathematikleistungen für die Analysen der vorliegenden Arbeit berücksichtigt werden sollten. „In der Literatur wird eine Vielzahl von Verursachungsfaktoren für die Rechenstörung diskutiert. Einigkeit scheint lediglich darin zu bestehen, dass mehrere Faktoren auf biologischer, (neuro-)psychologischer und psychosozialer Ebene sowie Umweltfaktoren die Dyskalkulie bedingen und diese Ursachen sich auch gegenseitig beeinflussen können“ (Thomas et al., 2015, S. 443). Die wichtigsten Faktoren für die Entstehung einer Rechenschwäche werden im Folgenden diskutiert.

2.3.1 Kognitive Faktoren

Ein grosser Teil der Untersuchungen zu Ursachenfaktoren für Rechenschwäche bezieht sich auf kognitive Faktoren, wobei sowohl domänenspezifische als auch domänenübergreifende Defizite als mögliche Ursache für eine Rechenschwäche diskutiert werden. Bei den domänenübergreifenden Faktoren stehen insbesondere die verschiedenen Komponenten des Arbeitsgedächtnisses im Fokus, bei den domänenspezifischen Defiziten wird hauptsächlich der Einfluss des sogenannten Zahlensinns („number sense“) diskutiert. Domänenübergreifend bedeutet hier, dass es sich um Fähigkeiten handelt, die sich auf verschiedene kognitive Prozesse und somit auf verschiedene kognitive Leistungen auswirken sollten. Mit domänenspezifischen Defiziten sind hingegen Fähigkeiten gemeint, die spezifisch für die mathematischen Leistungen relevant sind, sich aber nicht direkt auf andere kognitive Leistungen auswirken sollten. Im Folgenden werden zunächst Ergebnisse zum Einfluss des Arbeitsgedächtnisses bei Rechenschwäche berücksichtigt und in einem zweiten Schritt der Forschungsstand zu Beeinträchtigungen im Zahlensinn dargestellt.

Arbeitsgedächtnis

Studien, die sich mit dem Einfluss des Arbeitsgedächtnisses auf die mathematischen Leistungen befassen, beziehen sich in der Regel auf das *Arbeitsgedächtnismodell* von Baddeley (1986). Das Arbeitsgedächtnis ist wesentlich an der Verarbeitung von Informationen beteiligt und kann für kurze Zeit eine begrenzte Menge an Informationen speichern. In seiner ursprünglichen Konzeption ist das Arbeitsgedächtnis in drei Bereiche unterteilt: die zentrale Exekutive und die beiden Subsysteme der phonologischen Schleife und des visuell-räumlichen Skizzenblocks (Baddeley, 1986). Die *zentrale Exekutive* kontrolliert und koordiniert die beiden Subsysteme, steuert die Aufmerksamkeit und stellt die Verbindung zum Langzeitgedächtnis her. Die *phonologische Schleife* speichert und verarbeitet kurzfristig verbale und der *visuell-räumliche Skizzenblock* visuell-räumliche Informationen (Baddeley, 1986). Die zentrale Exekutive ist wichtig für die Steuerung von Rechenprozessen und somit auf prozeduraler Ebene von grosser Bedeutung. Die phonologische Schleife ist als verbaler Speicher bei Zählakten sowie beim Behalten von Ausgangsinformationen und Zwischenergebnissen von Rechnungen involviert (Baddeley, 1999). Der visuell-räumliche Skizzenblock hingegen ist wichtig, wenn es z.B. um die kurzzeitige Speicherung der Position einer Ziffer im Stellenwertsystem geht (Baddeley, 1999). Neuere Konzeptionen des Arbeitsgedächtnismodells (Baddeley, 1996; Miyake et al., 2000) unterteilen die zentrale Exekutive zusätzlich in weitere Unterbereiche: *Inhibition* (Unterdrückung nicht relevanter Informationen), *Shifting* (flexibler Aufgabenwechsel) und *Updating* (Aktualisierung des Arbeitsgedächtnisses).

In den vergangenen zwei Jahrzehnten haben sich diverse Studien mit der Frage beschäftigt, ob das Arbeitsgedächtnis generell einen Einfluss auf die mathematische Leistung hat und ob Kinder mit Rechenschwäche Defizite in diesem Bereich aufweisen, die an der Entstehung der mathematischen Schwierigkeiten beteiligt sein könnten (Friso-van den Bos, van der Ven, Kroesbergen & van Luit, 2013). Die dabei gewonnenen Erkenntnisse zum Einfluss des Arbeitsgedächtnisses sind jedoch sehr heterogen, sodass ein eindeutiges Fazit schwierig ist. In einigen Untersuchungen wurde ein eher generelles Defizit im Arbeitsgedächtnis bei rechenschwachen Kindern festgestellt (Geary et al., 2009; Mabbott & Bisanz, 2008; Szűcs, Devine, Soltész, Nobes & Gabriel, 2013; Toll, van der Ven, Kroesbergen & van Luit, 2011), es gibt aber auch Studien, die ein solches allgemeines Defizit nicht belegen konnten (z.B. Landerl, Bevan & Butterworth, 2004). Defizite in der zentralen Exekutive zeigten sich ebenfalls in mehreren Untersuchungen (Andersson, 2008b; Geary et al., 2007; McLean & Hitch, 1999; Wu et al., 2008), es gibt aber auch Studien, die keine Abweichungen bei rechenschwachen Kindern feststellen konnten (Schuchardt, Kunze, Grube & Hasselhorn, 2006; Vanbist et al., 2014), und solche, die nur bei Kindern mit einer kombinierten Lernschwäche (Rechenschwäche und Lese-/Rechtschreibschwäche) ein entsprechendes Defizit fanden (Schuchardt & Mähler, 2010). Ebenfalls nur bei einem Teil der Untersuchungen konnten Defizite bezüglich der phonologischen Schleife (Andersson, 2008b; Geary et al., 2007; Schuchardt et al., 2006; Swanson & Jerman, 2006; Vanbist et al., 2014) und des visuell-räumlichen Skizzenblocks (Andersson & Östergren, 2012; Geary et al., 2007; McLean & Hitch, 1999; Schuchardt & Mähler, 2010; Swanson & Jerman, 2006; Toll, Kroesbergen & van Luit, 2016) festgestellt werden.

Angesichts der grossen Heterogenität der Ergebnisse bietet die Metaanalyse von Friso-van den Bos et al. (2013) einige Hinweise zur Interpretation der uneinheitlichen Befunde. Dabei muss beachtet werden, dass diese Metaanalyse nicht nur Studien berücksichtigt, in denen rechenschwache Kinder untersucht wurden, sondern auch solche, an denen Kinder des gesamten Leistungsspektrums teilgenommen haben. In der Metaanalyse konnte festgestellt werden, dass alle Komponenten des Arbeitsgedächtnisses in den berücksichtigten Studien einen gewissen Einfluss auf die mathematische Leistung haben, Letzterer jedoch je nach Studie unterschiedlich stark ausfällt (Friso-van den Bos et al., 2013). Der grösste Einfluss zeigte sich dabei bei der exekutiven Funktion des verbalen Updating, gefolgt vom visuell-räumlichen Skizzenblock und vom visuell-räumlichen Updating (ebenfalls zentrale Exekutive). Etwas kleiner war der Einfluss der phonologischen Schleife, den geringsten Einfluss hatten die exekutiven Funktionen Inhibition und Shifting (Friso-van den Bos et al., 2013).

Die Gründe für die Uneinheitlichkeit der Ergebnisse sind vielfältig. So hat sich gezeigt, dass bei Stichproben mit älteren Kindern der Einfluss des Arbeitsgedächtnisses auf die mathematische Leistung abnimmt, sodass je nach Alter der untersuchten Kinder unterschiedliche Ergebnisse möglich sind (Friso-van den Bos et al., 2013). Auch die verwendeten Mathematiktests beeinflussen die Ergebnisse. Weiter hat sich gezeigt, dass eher curriculumsorientierte Tests mit höheren Korrelationen zwischen Arbeitsgedächtniskomponenten und Mathematikleistungen einhergehen als Tests, bei denen primär Rechenfertigkeiten oder isolierte mathematische Kompetenzen geprüft werden (Friso-van den Bos et al., 2013). Das wird damit begründet, dass beim eigentlichen Ausrechnen exekutive Funktionen eher weniger wichtig sind als z.B. beim Problemlösen, bei dem verschiedene Rechenschritte geplant und überwacht werden müssen. Auch die eingesetzten Aufgaben zur Erfassung der Arbeitsgedächtniskomponenten sind manchmal von Untersuchung zu Untersuchung unterschiedlich und beinhalten innerhalb der gleichen Komponente Aufgaben mit unterschiedlich komplexen Anforderungen (Friso-van den Bos et al., 2013). Dadurch wird die Vergleichbarkeit der Ergebnisse verschiedener Studien erschwert. Ebenfalls einen Einfluss auf die Ergebnisse hat die Auswahl der Stichprobengruppen, da sich für unterschiedliche Gruppen von Kindern (z.B. rechenschwach vs. nicht rechenschwach) unterschiedlich starke Korrelationen zwischen Arbeitsgedächtnis und Mathematikleistung zeigen (Friso-van den Bos et al., 2013). Zudem werden in den oben zitierten Studien teilweise auch sehr unterschiedliche Cut-off-Werte verwendet, wenn es um die Auswahl rechenschwacher Kinder geht. Trotz teilweise uneinheitlichen Ergebnissen kann als Zwischenfazit zum aktuellen Forschungsstand dennoch festgehalten werden, dass das Arbeitsgedächtnis bei einer Rechenschwäche eine wichtige Rolle spielt und verschiedene Komponenten des Arbeitsgedächtnisses einen Einfluss auf die Mathematikleistung haben. Für eine verstehensorientierte Mathematikförderung sollten daher Massnahmen getroffen werden, die das Arbeitsgedächtnis entlasten.

Zahlensinn

Eine weitere, vielfach diskutierte Ursachenkomponente bezieht sich auf Einschränkungen in domänenspezifischen Fähigkeiten. Damit sind Fähigkeiten gemeint, die spezifisch für mathematische Leistungen relevant sind. Die Ausgangslage für diese Position sind Forschungserkenntnisse, wonach Menschen von Geburt an über ein kognitives Kernsystem verfügen, das ihnen ermöglicht, Mengen und Numerositäten zu verarbeiten und zu vergleichen (Landerl & Kaufmann, 2008). Dehaene (1997) hat dafür den Begriff „number sense“ gewählt, was auf Deutsch als *Zahlensinn* bezeichnet wird (Landerl & Kaufmann, 2008). Innerhalb dieser Positionen sind jedoch wiederum verschiedene Hypothesen zu finden, wie ein eingeschränkter Zahlensinn zu Rechenschwierigkeiten führen kann (Andersson & Östergren, 2012; Vanbist et al., 2014). Zunächst wurde die Hypothese aufgestellt, dass der Zahlensinn bei Kindern mit Rechenschwäche per se defizitär sei, d.h., dass sie grundsätzlich Schwierigkeiten haben, Numerositäten zu verarbeiten und zu vergleichen (Butterworth, 2005). Eine empirische Bestätigung für entsprechende Schwierigkeiten konnte mithilfe von Mengenvergleichsaufgaben bei rechenschwachen Kindern gegeben werden (z.B. de Smedt, Noël, Gilmore & Ansari, 2013). Dabei muss jedoch zwischen symbolischen und nicht symbolischen Mengenvergleichsaufgaben unterschieden werden, wie im Folgenden aufgezeigt wird. Bei nicht symbolischen Mengenvergleichen müssen zwei Bilder mit unterschiedlicher Anzahl Punkten verglichen werden, und es muss entschieden werden, welches Bild mehr Punkte enthält (de Smedt et al., 2013).[5] Bei symbolischen Mengenvergleichen werden hingegen direkt zwei Zahlen in arabischer Schreibweise präsentiert, und es muss entschieden werden, welche der beiden Zahlen grösser ist.

In einer Übersichtsarbeit haben Noël und Rousselle (2011) festgestellt, dass Kinder mit Rechenschwäche zwar in diversen Studien signifikant schlechtere Leistungen im symbolischen Mengenvergleich zeigten, Schwierigkeiten im nicht symbolischen Mengenvergleich jedoch nur ganz vereinzelt empirisch bestätigt wurden.[6] Demnach ist eher nicht anzunehmen, dass Kinder mit Rechenschwäche von Geburt an Probleme mit Bezug auf das approximative Zahlensystem aufweisen und somit nicht per se einen eingeschränkten Zahlensinn haben. Die beiden Autorinnen haben daher die Hypothese aufgestellt, dass nicht der Zahlensinn an sich beeinträchtigt ist, sondern Kinder mit Rechenschwäche Probleme beim Zugriff auf die semantische Grössenrepräsentation von Zahlen haben und es sich somit um spezifisch numerische Schwierigkeiten mit Bezug auf Mengenvergleiche handelt (Noël & Rousselle, 2011). Auch weitere Studien bzw. Übersichtsarbeiten (z.B. de Smedt & Gilmore, 2011; de Smedt et al., 2013; Vanbist et al., 2014) bestätigen die Hypothese von Noël und Rousselle (2011), wonach rechenschwachen Kindern der Zugriff auf die Zahlensemantik Schwierigkeiten bereitet, nicht jedoch die Hypothese von Butterworth (2005) u.a., wonach bei Rechenschwäche der Zahlensinn per se beeinträchtigt ist. Allerdings gibt es auch Studien, die Schwierigkeiten bei nicht symbolischen Mengenvergleichen belegen konnten (Mussolin, Mejias & Noël, 2010; Piazza et al., 2010). Dazu muss berücksichtigt werden, dass auch hier methodische Unterschiede (z.B. Aufgabendesign, Mass für die Messung der Leistung) zu den unterschiedlichen Ergebnissen beigetragen haben können (de Smedt et al., 2013).

Interessante Hinweise zum Einfluss des Zahlensinns bei Rechenschwäche bietet auch die neuropsychologische Forschung. Diese stützt sich auf das Triple-Code-Modell von Dehaene (1992) und Dehaene und Cohen (1995), wonach sich der Zahlensinn in Form der analogen Grössenrepräsentation hirnanatomisch dem intraparietalen Sulcus (IPS) zuordnen lässt (vgl. Kap. 2.2.1). Diverse Studien belegen, dass der IPS eine wichtige Rolle bei der Verarbeitung von Zahlenmengen bei gesunden erwachsenen Personen spielt (Ansari, Garcia, Lucas, Hamon & Dhital, 2005; Kaufmann et al., 2006; Pinel et al., 1999; Pinel, Dehaene, Rivière & LeBihan, 2001). Zudem konnte gezeigt werden, dass bei Kindern mit einer Rechenschwäche bei numeri-

[5] Meistens wird hierfür sichergestellt, dass die Punkte auf den beiden Bildern trotz unterschiedlicher Anzahl insgesamt die gleiche Fläche einnehmen, sodass nicht die Fläche als Basis für den Mengenvergleich herangezogen wird. Auch andere Faktoren wie die Dichte der Punkte werden entsprechend kontrolliert.
[6] Die schlechteren Leistungen sind dabei meistens nicht in den Fehlerquoten selbst zu finden, sondern zeigen sich in längeren Reaktionszeiten oder anderen Messungen (z.B. Distanzeffekt).

schen Schätzaufgaben ebenso wie bei nicht symbolischen Mengenvergleichen der IPS schwächer aktiviert wird als bei Kindern ohne Rechenschwäche, was als atypische funktionelle Aktivierung des IPS bezeichnet wird (Kucian et al., 2006; Price, Holloway, Rasanen, Vesterinen & Ansari, 2007). Allerdings gibt es auch Untersuchungen, die eine erhöhte Aktivierung des IPS bei Kindern mit einer Rechenschwäche belegen konnten (Kaufmann, Vogel, Starkey, Kremser & Schocke, 2009; Kucian, Loenneker, Martin & von Aster, 2011). Aufgrund vergleichbarer Fehlerquoten und Reaktionszeiten bei Kindern mit und ohne Rechenschwäche wird die erhöhte Aktivierung des IPS jedoch als kompensatorischer Mechanismus interpretiert, sodass auch hier von einer atypischen funktionellen Aktivierung des IPS gesprochen wird (Kucian, Loenneker et al., 2011). Zu erwähnen ist diesbezüglich, dass die Wirkungsrichtung noch weitestgehend ungeklärt ist. D.h., es kann derzeit nicht gesagt werden, ob die atypische funktionelle Aktivierung des IPS eine Folge einer Rechenschwäche oder ein Verursachungsfaktor für eine Rechenschwäche ist (Vogel & Ansari, 2012). Somit weisen auch neuropsychologische Studien darauf hin, dass rechenschwache Kinder Schwierigkeiten in Zusammenhang mit dem Zahlensinn haben. Für eine mathematische Förderung muss daher beachtet werden, dass die Förderung des Zahlensinns wichtig ist.

Abschliessend kann festgehalten werden, dass neben domänenübergreifenden Aspekten (Arbeitsgedächtnis) bei rechenschwachen Kindern auch domänenspezifische Faktoren (Zahlensinn) eine zentrale Rolle zu spielen scheinen und somit als Ursachenfaktoren für eine Rechenschwäche infrage kommen. In der Untersuchung von Toll et al. (2016) konnten diese beiden Faktoren jedoch nur 24% (mathematisches Faktenwissen) bzw. 28% (Sachaufgaben) der Varianz in den Mathematikleistungen erklären. Es muss daher angenommen werden, dass zusätzlich zu den bisher dargestellten domänenspezifischen und domänenübergreifenden kognitiven Faktoren weitere Aspekte ebenfalls einen bedeutsamen Einfluss auf die Entstehung einer Rechenschwäche haben. Einige dieser Aspekte werden im Folgenden diskutiert.

2.3.2 Intelligenz

Intelligenz ist ein wichtiger individueller Faktor, der sich auf Schulleistungen im Allgemeinen auswirkt, wobei es verschiedene Intelligenzdefinitionen gibt, die aber einen gemeinsamen Kern aufweisen: „Stellvertretend für alle Intelligenzdefinitionen kann die von Wechsler (1964) stehen, wonach Intelligenz die zusammengesetzte oder globale Fähigkeit des Individuums ist, die ihm erlaubt, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinander zu setzten“ (Breitenbach, 2016, S. 359). Als Beispiel soll hier die Intelligenztheorie von Cattell (1965) erwähnt werden, in der zwei sogenannte g-Faktoren angenommen werden. Die fluide Intelligenz wird darin als angeborene und eher allgemeine Fähigkeit verstanden, sich mit neuartigen Problemen auseinanderzusetzen, die kristalline Intelligenz hingegen baut auf erworbenem Wissen auf und ist somit durch Lern- und Bildungsprozesse beeinflusst (Cattell, 1965). Weitere Intelligenztheorien sollen an dieser Stelle nicht ausgeführt werden.

Grundsätzlich kann festgehalten werden, dass die Intelligenz einen Einfluss auf Schulleistungen im Allgemeinen ebenso wie auf Mathematikleistungen im Speziellen hat (W. Schneider et al., 2013). Grube und Hasselhorn (2006, S. 88) kommen aufgrund des Forschungsstandes zum Schluss, dass „circa 25 bis 45% der Varianz von Schulleistungen […] sich durch Intelligenzleistungen aufklären [lassen]“. In einer eigenen Untersuchung konnten sie jedoch zeigen, dass die nonverbale Intelligenz zwar mit Mathematikleistungen korreliert, das Vorwissen jedoch der bedeutendste Prädiktor für Mathematikleistungen und für Schulleistungen im Allgemeinen ist (Grube & Hasselhorn, 2006). Dieser Befund wird so interpretiert, dass die Intelligenz einen indirekten Einfluss auf die Mathematikleistung hat, indem die Intelligenz den Erwerb von Vorwissen und das Vorwissen wiederum die aktuelle Mathematikleistung beeinflusst (Grube & Hasselhorn, 2006). Auch aufgrund von Ergebnissen aus der LOGIK-Studie können vergleichbare Schlüsse gezogen werden:

> Der Einfluss der Intelligenz macht sich vorwiegend in der sogenannten konfundierten Varianz bemerkbar. Diese sagt, dass sich Kinder mit einer höheren Intelligenz auf Dauer mehr mathematisches Wissen aneignen und deshalb bessere Leistungen erbringen. Der durch Intelligenzunterschiede vermittelte Varianzanteil des mathematischen Vorwissens ist jedoch deutlich geringer als der von der Intelligenz unabhängige Varianzanteil. (Stern, 2009, S. 162)

Mit Blick auf Kinder mit Rechenschwäche, die im Vergleich zu Schülerinnen und Schülern ohne entsprechende Schwierigkeiten über weniger tragfähiges Vorwissen verfügen, gibt es jedoch Hinweise, wonach die Intelligenz zusätzlich zum mathematischen Vorwissen ebenfalls einen bedeutsamen Einfluss auf die mathematische Leistung hat (Moser Opitz et al., 2017). Dabei zeigte sich aber auch für Kinder mit Rechenschwäche, dass der Einfluss der Intelligenz weniger gross ist als der des Vorwissens (Moser Opitz et al., 2017). Daher soll im Rahmen der vorliegenden Arbeit der Frage nachgegangen werden, welchen Einfluss die Intelligenz unter gleichzeitiger Berücksichtigung des mathematischen Vorwissens in der hier untersuchten Stichprobe hat, und geprüft werden, ob sich dabei auch Klassenkompositionseffekte zeigen.

2.3.3 Sprachliche Kompetenzen

Bereits seit längerem wird im schulischen Kontext diskutiert, dass sprachliche Kompetenzen einen bedeutsamen Einfluss auf die Mathematikleistung und insbesondere auf das Lösen von Text- und Sachaufgaben haben (z.B. Moser Opitz, 2013). So konnten sprachliche Kompetenzen für den englischsprachigen Raum als wichtiger Prädiktor für die Mathematikleistung bestätigt werden (z.B. LeFevre et al., 2010). Auch für den deutschsprachigen Raum wurde ein signifikanter Zusammenhang zwischen Lesekompetenz bzw. Rechtschreibkompetenz und Mathematikleistung sowie den mathematischen Fortschritten eines Kindes gefunden (z.B. Nikolova, 2011; Paetsch, Radmann, Felbrich, Lehmann & Stanat, 2016). Dieser Sachverhalt trifft besonders auch auf Kinder zu, die eine andere Familiensprache sprechen als die Instruktionssprache in der Schule (Paetsch et al., 2016). Im Rahmen der TIMSS-Studie konnte für die Erhebungen für das Jahr 2011 gezeigt werden, dass Zweitsprachenlernende auch unter Kontrolle des sozioökonomischen Status über schlechtere Mathematikleistungen verfügen als Kinder, die zu Hause fast immer oder immer Deutsch sprechen (Tarelli, Schwippert & Stubbe, 2012). Im internationalen Vergleich belegt dieselbe TIMSS-Studie zudem, dass nur in einer Minderheit der beteiligten Länder keine signifikanten Leistungsunterschiede bestehen zwischen Kindern, die zu Hause regelmässig die Landessprache sprechen, und Kindern, die dies fast oder gar nie tun (Tarelli et al., 2012).

Bei Kindern mit Deutsch als Zweitsprache, die niedrigere sprachliche Kompetenzen aufweisen, konnte für Deutschland auch belegt werden, dass ein Zusammenhang zwischen mathematischem Fachwortschatz und allgemeinen Wortschatzkenntnissen sowie Grammatikkompetenz besteht (Paetsch, Felbrich & Stanat, 2015). Hierzu muss beachtet werden, dass fachsprachliche Kompetenzen nicht per se gegeben sind, sondern von der Lehrperson vermittelt werden müssen, damit sich die Schülerinnen und Schüler diese aktiv aneignen können (Moser Opitz & Schindler, 2017). Die Versprachlichung mathematischer Handlungen im Allgemeinen wird zudem gerade für rechenschwache Kinder als wichtig für die mathematische Entwicklung und den Aufbau von konzeptuellem Verständnis angesehen (Scherer & Moser Opitz, 2010; Werner, 2009). Paetsch et al. (2015, S. 27) kommen zum Schluss, „[…] dass sprachliche Fähigkeiten nicht nur zur Bewältigung der sprachlichen Anforderungen der konkreten Aufgaben, sondern auch beim Erwerb und Abruf des mathematischen Wissens eine wichtige Rolle für Zweitsprachenlernende spielen“. Ein indirekter Einfluss der sprachlichen Kompetenzen wird auch von Vukovic und Lesaux (2013) bestätigt, wonach Zweitsprachenlernende grössere Mühe haben, sich zentrale mathematische Konzepte und Repräsentationsformen in gleicher Weise anzueignen, wie Kinder, deren Muttersprache der Instruktionssprache entspricht. Prediger, Wilhelm, Büchter, Gürsoy und Benholz (2015) konnten zeigen, dass sich bei Schülerinnen und Schülern mit sprachlichen Schwierigkeiten beim Lösen von Textaufgaben insbesondere prozessuale und konzeptuelle Hürden stellen.

Ein weiterer Aspekt, der im Zusammenhang mit sprachlichen Anforderungen immer wieder betont wird, ist die Schwierigkeit der Zahlwortproduktion im Deutschen, was sich auf die Zählentwicklung und alle darauf aufbauenden mathematischen Kompetenzen auswirken kann (Moser Opitz, 2013). Eine Herausforderung in der deutschen Sprache ist, dass bei zweistelligen Zahlen zuerst der Einer und erst danach der Zehner ausgesprochen werden muss (Zehner-Einer-Inversion, z.B. einundzwanzig), zudem müssen z.B. im Vergleich zu gewissen asiatischen Sprachen im Deutschen mehr Zahlwörter auswendig gelernt werden (Moser Opitz, 2013). So können die Zahlwörter für die Zahlen eins bis zwölf nicht abgeleitet werden, zudem muss die Bildung der Zahlwörter für die Zehnerzahlen erlernt werden (Schäfer, 2005). Dass bei Zahlen ab dreizehn zuerst der Einer und dann der Zehner (mit Ausnahme der Zehnerzahlen) und ab der Zahl hundert zusätzlich zuerst der Hunderter, dann der Einer und erst zum Schluss der Zehner genannt werden muss, erschwert zusätzlich die Einsicht in das Notationsystem der Zahlen im Dezimalsystem (Schäfer, 2005; W. Schneider et al., 2013).

Bezüglich sprachlicher Aspekte muss auch beachtet werden, dass der Anteil Kinder, die sowohl eine Rechenschwäche als auch eine Lese-/Rechtschreibschwäche aufweisen, grösser ist als ursprünglich angenommen (Fischbach et al., 2013; von Aster et al., 2007). Forschungsergebnisse für Kinder mit einer kombinierten Lernschwäche zeigen, dass diese nicht grundsätzlich über qualitativ unterschiedliche Schwierigkeiten verfügen im Vergleich zu Kindern, die eine isolierte Rechenschwäche haben (Andersson, 2010; Jordan & Hanich, 2003; Landerl et al., 2004). Allerdings sind auch Belege vorhanden, wonach in gewissen Inhaltsbereichen die Schwierigkeiten von Kindern mit einer kombinierten Lernschwäche grösser sind als bei Kindern mit einer isolierten Rechenschwäche, wobei diesbezüglich insbesondere Schwierigkeiten bei Aufgaben, die Problemlösekompetenzen erfassen, zu nennen sind (Fuchs & Fuchs, 2002; Geary, Hoard & Hamson, 1999; Jordan & Hanich, 2003; Landerl et al., 2004; Powell, Fuchs, Fuchs, Cirino & Fletcher, 2009a). Zudem gibt es Hinweise, dass Kinder mit einer kombinierten Lernschwäche weniger stark von mathematischen Förderprogrammen profitieren als Kinder mit einer isolierten Rechenschwäche (Fuchs, Fuchs & Prentice, 2004; Powell, Fuchs, Fuchs, Cirino & Fletcher, 2009b). Ebenfalls zu beachten ist, dass mit einer kombinierten Lernschwäche gravierendere Defizite im Arbeitsgedächtnis einhergehen können als bei einer isolierten Rechenschwäche (Peng & Fuchs, 2016).

Zusammenfassend kann festgehalten werden, dass sprachliche Kompetenzen einen Einfluss auf die mathematische Entwicklung haben, dass sich diese Problematik für Zweitsprachenlernende verstärkt zeigen kann und fachsprachliche Kompetenzen aktiv angeeignet werden müssen. Der Versprachlichung mathematischer Handlungen kommt daher einen wichtigen Stellenwert in einer Mathematikförderung zu und der Einfluss sprachlicher Faktoren sollte auch in mathematischen Interventionsstudien berücksichtigt werden. Die deutsche Sprache ist hinsichtlich der Zahlwortproduktion zudem gerade für fremdsprachige Kinder herausfordernd. Des Weiteren muss davon ausgegangen werden, dass bei Kindern, die zusätzlich zu einer Rechenschwäche auch eine Lese-/Rechtschreibschwäche aufweisen, sich inhaltlich zwar vergleichbare Schwierigkeiten zeigen, teilweise aber auch grössere Leistungsrückstände anzunehmen sind.

2.3.4 Geschlecht

Die Frage nach Unterschieden in den mathematischen Leistungen zwischen den Geschlechtern findet sich in vielen Büchern und Artikeln zum Thema und wird auch mit Bezug auf schulische Leistungen in anderen Fächern (z.B. Naturwissenschaften, Sprache) rege diskutiert. Bezüglich Geschlechterunterschiede in mathematischen Leistungen sind unterschiedliche Ergebnisse vorhanden. Eine ältere, englischsprachige Metastudie konnte über alle einbezogenen Stichproben hinweg bessere Mathematikleistungen beim männlichen Geschlecht feststellen, wobei die durchschnittliche Effektstärke nicht als sehr hoch eingeschätzt wird (Effektstärke $d = 0.15$; Hyde, Fennema & Lamon, 1990). Eine neuere Metaanalyse, die englischsprachige Publikationen seit den 1990er-Jahren einbezogen hat, konnte jedoch keine Geschlechterunterschiede mehr finden ($d = 0.05$; Lindberg, Hyde, Petersen & Linn, 2010). Diesbezüglich muss allerdings beachtet werden, dass es durchaus Unterschiede zwischen einzelnen Ländern gibt. Im Rahmen der PISA-2000-Untersuchung konnte für die Hälfte der Länder ein Geschlechtereffekt mit Vorteil für die Knaben festgestellt werden, was auch auf die Schweiz zutraf (Bundesamt für Statistik & Schweizerische Konferenz der kantonalen Erziehungsdirektoren, 2002). Der Unterschied wird jedoch als wenig gravierend eingeschätzt ($d = 0.14$; S. 42). Auch in den PISA-2003-Erhebungen wurde für die Schweiz ein signifikanter Geschlechterunterschied zugunsten der Knaben ermittelt (Bundesamt für Statistik & Schweizerische Konferenz der kantonalen Erziehungsdirektoren, 2004).

Bezüglich der Häufigkeit von Rechenstörungen nach Geschlecht ist die Forschungslage ebenfalls sehr heterogen. Gemäss Fischbach et al. (2013) gibt es empirische Studien, die keinen Geschlechterunterschied für die Häufigkeit von Rechenstörungen finden konnten (z.B. Desoete, Roeyers & DeClercq, 2004), solche, die höhere Prävalenzraten bei Mädchen festgestellt haben (z.B. Hein, Bzufka & Neumärker, 2000), und solche, bei denen Knaben übervertreten waren (z.B. Reigosa-Crespo et al., 2012). Für die Schweiz liegt eine Studie vor, die für isolierte Rechenstörungen keine Geschlechtsunterschiede feststellen konnte, für kombinierte Rechen- und Lese-/Rechtschreibstörung jedoch eine leicht höhere Prävalenzrate für Knaben belegt (von Aster et al., 2007).

Zwar werden durchaus auch biologische Faktoren für entsprechende Geschlechterunterschiede diskutiert (siehe Krinzinger & Günther, 2013). Soziokulturelle Faktoren werden derzeit jedoch eher als ausschlaggebend betrachtet. So zeigte eine Analyse der PISA-Daten des Jahres 2003, dass Masse für die Gleichstellung der Geschlechter (z.B. prozentualer Anteil von Frauen in der Forschung) in den untersuchten Ländern signifikante Prädiktoren für die gefundenen Geschlechterunterschiede in den Mathematikleistungen sind (Else-Quest, Hyde & Linn, 2010). In dieser Untersuchung wurden auch Einstellungen im Fach Mathematik aus TIMSS und PISA analysiert, und es zeigte sich, dass Knaben im Vergleich zu Mädchen ein höheres Selbstvertrauen in Mathematik aufweisen, positivere Einstellungen zum Fach haben und von weniger Mathematikangst berichten (Else-Quest et al., 2010). Alle diese Faktoren korrelierten signifikant mit der Mathematikleistung. Deshalb kann angenommen werden, dass Geschlechterunterschiede mit Bezug auf Einstellungen zum Fach Mathematik ebenfalls durch gesellschaftliche Rollenverteilungen und -erwartungen und somit ebenso soziokulturell bedingt sind (Else-Quest et al., 2010). In diesem Zusammenhang wird zudem diskutiert, dass auch Lehrpersonen geschlechterstereotype Erwartungen mit Bezug auf die schulischen Leistungen von Kindern haben können (siehe z.B. Moser Opitz, 2013). So konnte eine Studie belegen, dass unter Kontrolle des Lern- und Unterrichtverhaltens (z.B. Bereitschaft, sich anzustrengen) die gleichen mathematischen Leistungen von Knaben im Vergleich zu Mädchen durch die Lehrpersonen höher bewertet wurden, was dazu beitrug, dass eine zunehmende Geschlechterdifferenz in den mathematischen Leistungen beobachtet werden konnte (Robinson-Cimpian, Theule Lubienski, Ganley & Copur-Gencturk, 2014).

Zusammenfassend kann daher festgehalten werden, dass sich in einigen, aber nicht in allen Untersuchungen Geschlechterunterschiede mit Bezug auf die Mathematikleistung zeigen, dass diese jedoch eher durch soziokulturelle Faktoren erklärt wird und nicht biologische Gründe für diese Unterschiede angenommen werden. Inwiefern Geschlechtereffekte in der in dieser Arbeit zu untersuchenden Stichprobe festzustellen sind, soll daher überprüft werden.

2.3.5 Sozioökonomischer Status

Sowohl für den deutschsprachigen Raum als auch im internationalen Kontext konnte in verschiedenen Studien belegt werden, dass der sozioökonomische Status (SES) einen Einfluss auf Schulleistungen im Allgemeinen und auf die Mathematikleistungen im Speziellen hat (vgl. Jordan & Levine, 2009; Stubbe, Tarelli & Wendt, 2012). Dabei muss beachtet werden, dass es verschiedene Operationalisierungen für den sozioökonomischen Status gibt, die in der Forschung verwendet werden. Zum grössten Teil beziehen sich diese auf die Arbeit von Bourdieu (1983) und seine Unterscheidung zwischen ökonomischem, kulturellem und sozialem Kapital (Stubbe et al., 2012). Nebst der Bücheraufgabe (vgl. Kap. 5.4.3) wird der SES daher meist über das Bildungsniveau, den Berufsstatus oder das Einkommen der Eltern erfasst (Stubbe et al., 2012).

Für eine deutsche Stichprobe konnten Krajewski und Schneider (2006) zeigen, dass der SES operationalisiert über den elterlichen Schulabschluss und Beruf einen signifikanten Einfluss auf die Mathematikleistung am Ende der vierten Klasse hat, sodass Schülerinnen und Schüler mit geringerem SES durchschnittlich schlechtere Leistungen aufweisen. Für die erste Klasse zeigten sich in dieser Studie jedoch noch keine Effekte mit Bezug auf mathematische Vorläuferfertigkeiten. In einer anderen Studie, in der der SES über das Bildungsniveau erfasst wurde, konnte jedoch auch für mathematische Vorläuferfertigkeiten ein signifikanter Einfluss des SES festgestellt werden (Schuchardt, Piekny, Grube & Mähler, 2014). Für die frühen Schuljahre muss jedoch von uneinheitlichen Ergebnissen ausgegangen werden, da auch Dummert, Endlich, Schneider und Schwenck (2014) für die Mathematikleistung in der vierten Klasse, nicht jedoch für die zweite Klasse einen signifikanten Einfluss des SES, erfasst über den Beruf und das Einkommen der Eltern, belegen konnten. Auch im Rahmen der TIMSS-Untersuchungen zeigt sich für die vierte Klasse ein signifikanter Effekt des SES für die Mathematikleistung, wobei dort die Bücheraufgabe eingesetzt und in Form einer Dummy-Codierung in die Analysen aufgenommen wurde (mehr vs. weniger als 100 Bücher; Stubbe et al., 2012).

Für die Erklärung des Einflusses des SES auf die Mathematikleistung werden insbesondere das elterliche Unterstützungsverhalten, elterliche Bildungsaspirationen und familiäre Lerngelegenheiten als wichtig erachtet (Jordan & Levine, 2009; Krajewski & Schneider, 2006). Entsprechend wird angenommen, dass sich das mathematische Vorwissen bereits zum Zeitpunkt der Einschulung aufgrund mehr oder weniger förderlicher Lern- und Entwicklungsbedingungen im Elternhaus unterscheiden kann (Jordan & Levine, 2009). In diesem Zusammenhang ist auch zu beachten, dass ein Teil der Varianz, der sich durch den SES erklären lässt, sich mit anderen

Einflussfaktoren überschneiden kann, wobei diesbezüglich z.B. der Migrationshintergrund und sprachliche Leistungen im Bereich der Schulsprache zu nennen sind (Dummert et al., 2014; Jordan & Levine, 2009). Ein Grund hierfür ist, dass Schülerinnen und Schüler mit Migrationshintergrund überproportional häufig einen niedrigeren SES aufweisen (Jordan & Levine, 2009). Aufgrund der genannten Faktenlage soll daher unter Berücksichtigung weiterer beeinflussender Variablen untersucht werden, welchen Einfluss der SES auf die mathematischen Leistungen hat.

2.3.6 Motivationale und emotionale Faktoren

Forschungsergebnisse belegen, dass auch emotionale und motivationale Aspekte bei Rechenschwäche eine Rolle spielen. So zeigen Untersuchungen, dass rechenschwache Schülerinnen und Schüler stärker von Mathematikangst betroffen sind als Schülerinnen und Schüler ohne mathematische Schwierigkeiten (z.B. Lebens, Graff & Mayer, 2011; Sorvo et al., 2017). Ebenso weisen rechenschwache Schülerinnen und Schüler ein negativeres fachbezogenes, akademisches Selbstkonzept auf (z.B. Schuchardt et al., 2015) und haben negativere Einstellungen zum Fach Mathematik (z.B. Jahnke-Klein, 2001; Moser Opitz, 2013). Zudem kann sich die Mathematikangst auch in Form einer grösseren Prüfungsangst zeigen, was wiederum Leistungen in Prüfungen negativ beeinflussen kann (Pixner & Kaufmann, 2013).

Mit Blick auf die emotionalen und motivationalen Faktoren wird angenommen, dass negative Lernerfahrungen im Fach Mathematik dazu beitragen, dass rechenschwache Schülerinnen und Schüler sich als wenig selbstwirksam erleben und mathematikbezogene Ängste sowie eine negativere Einstellung zum Fach entwickeln, was die Bereitschaft, sich anzustrengen bzw. sich mit komplexen mathematischen Sachverhalten auseinanderzusetzen, negativ beeinflusst (Huck & Schröder, 2016; Kohn et al., 2013). Dies wirkt sich wiederum negativ auf die weitere mathematische Entwicklung aus, sodass ein regelrechter Teufelskreis entstehen kann. In diesem Sinne können die hier genannten Faktoren sowohl als Folge wie auch als Ursache einer Rechenschwäche angesehen werden. Ebenfalls zu beachten ist in diesem Zusammenhang, dass dabei auch geschlechterspezifische Effekte anzunehmen sind. So zeigen Studien, dass Mädchen ein negativeres mathematisches Selbstkonzept aufweisen und negativere Einstellungen zum Fach Mathematik haben (Jahnke-Klein, 2001; Manger & Eikeland, 1998). Es kann daher als wichtig erachtet werden, dass rechenschwachen Schülerinnen und Schülern ein anspruchsvolles Lernen auf ihrem Niveau ermöglicht wird, damit sie sich als selbstwirksam erleben und positive emotionale sowie motivationale Assoziationen zum Fach Mathematik aufbauen können.

Emotionale und motivationale Aspekte sollen daher für die Gestaltung der unterrichtsintegrierten Mathematikförderung berücksichtigt werden.

2.3.7 Unterricht

Dass der Mathematikunterricht ebenfalls einen Einfluss auf die Entstehung einer Rechenschwäche haben kann, gilt heute als unbestritten (z.B. Gold, 2016; Landerl & Kaufmann, 2008; Moser Opitz, 2013; Scherer & Moser Opitz, 2010). Dabei geht es einerseits um die Frage, wie Mathematikunterricht qualitativ gestaltet wird, wobei insbesondere inhaltliche, didaktische und mathematikdidaktische Aspekte des Unterrichtsangebotes als relevant erachtet werden (Scherer & Moser Opitz, 2010). Andererseits spielen aber auch quantitative Aspekte der Zeitnutzung eine Rolle, d.h. ob die zur Verfügung stehende Zeit „[…] in hohem Maße für fachliche Lernaktivitäten genutzt wird und […] es gelingt, Schülerinnen und Schüler im Unterricht mit nachweislich lernwirksamen Aktivitäten zu beschäftigen“ (Brühwiler, Helmke & Schrader, 2017, S. 302f).

Auf qualitativer Ebene müssen erstens zentrale Inhalte auf Basis von curricularen Vorgaben, mathematikdidaktischen Überlegungen und empirischen Erkenntnissen bestimmt werden, die im Unterricht behandelt werden sollen (Krauthausen & Scherer, 2007). Dabei sind entsprechende Erkenntnisse zu spezifischen Schwierigkeiten rechenschwacher Schülerinnen und Schüler speziell zu berücksichtigen (Scherer & Moser Opitz, 2010; W. Schneider et al., 2013). Eine ausführliche Darstellung inhaltlicher Schwierigkeiten rechenschwacher Lernender findet sich im Kapitel 2.4 der vorliegenden Arbeit. Zweitens werden auf didaktischer Ebene zusätzlich zur Klassenführung als zentrale Rahmenbedingung für einen qualitativ hochstehenden Unterricht folgende Aspekte als wesentlich erachtet: „Lernförderliches Unterrichtsklima; vielfältige Motivierung; Strukturiertheit, Klarheit, Verständlichkeit; Ziel-, Wirkungs- und Kompetenzorientierung; Schülerorientierung, Unterstützung; Aktivierung, Förderung selbstständigen Lernens; angemessene Variation von Methoden und Sozialformen; Konsolidierung, Sicherung, intelligentes Üben sowie Passung und sensibler Umgang mit heterogenen Lernvoraussetzungen“ (Brühwiler et al., 2017, S. 303; vgl. auch Helmke, 2010). Einige dieser Aspekte werden in der vorliegenden Arbeit ebenfalls ausführlich thematisiert (vgl. Kap. 3.2 und Kap. 3.3.1). Zusätzlich müssen drittens immer auch mathematikdidaktische Überlegungen einbezogen werden, um einen qualitativ hochstehenden Mathematikunterricht zu ermöglichen (vgl. Kap. 3.3). Damit Lehrpersonen einen qualitativ und quantitativ hochstehenden Mathematikunterricht anbieten können, sind sie somit in verschiedenen Bereichen auf professionelle Kompetenzen und auf professionelles Wissen angewiesen (vgl. Kap. 3.4).

Häufig geäusserte Kritikpunkte hinsichtlich der Gestaltung von Mathematikunterricht beziehen sich darauf, dass in vielen Fällen zu stark auf die Beherrschung von Rechenfertigkeiten fokussiert werde, Schulbücher primär im Sinne von Pflichtstoff abgearbeitet würden und der Aufbau von Verständnis vernachlässigt werde (Ginsburg, 1997; Landerl & Kaufmann, 2008; Moser Opitz, 2013; Wember, 2009). Ebenfalls als problematisch wird erachtet, wenn bei schwachen Lernenden Lernstoff zu stark reduziert wird, Zahlenräume künstlich beschränkt werden, nur sehr kleinschrittig vorgegangen wird, fixe Lösungswege vorgegeben werden und grössere Zusammenhänge in Einzelfakte zerlegt werden, um Überforderung möglichst auszuschliessen (Scherer, 1995). Zudem orientiert sich der Mathematikunterricht an Regelschulen häufig an durchschnittlich leistungsfähigen Schülerinnen und Schülern und berücksichtigt nur bedingt, wie auch leistungsschwache Lernende Einsicht und Verständnis aufbauen können (Lambert, 2015). Die genannten Kritikpunkte sind daher für die Gestaltung einer mathematischen Förderung ebenfalls zu berücksichtigen und der Verstehensorientierung sollte ein hoher Stellenwert zukommen.

2.4 Mathematische Schwierigkeiten rechenschwacher Kinder

In den letzten Jahren wurde die Forschung zu spezifischen mathematischen Schwierigkeiten von Kindern mit Rechenschwäche intensiviert. Dabei zeigte sich, dass rechenschwache Lernende bei bestimmten mathematischen Inhalten und Kompetenzen grössere Schwierigkeiten aufweisen als Kinder ohne Rechenschwäche. Allerdings kann nicht davon ausgegangen werden, dass sich die mathematischen Lernprozesse von rechenschwachen Kindern qualitativ von jenen von Kindern ohne Rechenschwäche unterscheiden (Lorenz & Radatz, 1993). So haben auch Kinder ohne Rechenschwäche beim Erwerb von spezifischen Inhalten gewisse Schwierigkeiten (vgl. z.B. Kap. 2.4.3). Im Folgenden werden die einzelnen Inhaltsbereiche jeweils separat thematisiert, und es wird der jeweilige Stand der empirischen Forschung zu den inhaltlichen Schwierigkeiten rechenschwacher Schülerinnen und Schüler aufgezeigt. Damit sollen mathematische Inhalte und Kompetenzen identifiziert werden, die für die inhaltliche Ausgestaltung einer unterrichtsintegrierten und verstehensorientierten Mathematikförderung berücksichtigt werden sollten. Da in dieser Arbeit eine Mathematikförderung in der dritten Klasse der Primarschule analysiert wird, stehen mathematische Vorläuferfertigkeiten, die sich in der Regel bereits vor Schulbeginn entwickeln, nicht im Zentrum des Interesses. Dennoch sollen diese im Folgenden kurz thematisiert werden, da sie eine zentrale Basis für die weitere mathematische Entwicklung darstellen (vgl. Kap. 2.2).

2.4.1 Vorläuferfertigkeiten

Untersuchungen im Kindergarten zeigen, dass rechenschwache Kinder bereits in diesem Alter in wichtigen mathematischen Vorläuferfertigkeiten Schwierigkeiten aufweisen und dass diese die späteren mathematischen Leistungen in der ersten und der vierten Klasse vorwegnehmen können (Krajewski, 2003; Krajewski & Schneider, 2006, 2009). Gemäss dem Entwicklungsmodell von Krajewski (2003, 2005, 2007) können unter Vorläuferfertigkeiten Kompetenzen verstanden werden, die in ihrem Modell auf Ebene 1 und Ebene 2 verortet werden (Krajewski & Schneider, 2006, 2009; vgl. Abbildung 3). So schnitten rechenschwache Kindergartenkinder signifikant schlechter ab bei Aufgaben, die numerische Basisfertigkeiten und das einfache Zahlenverständnis prüfen (z.B. Beherrschen der Zahlwortreihe, Grössenrepräsentationen im Sinne der Mächtigkeit von Zahlen; Krajewski & Schneider, 2009). Die Bedeutung von mathematischen Vorläuferfertigkeiten konnte auch in Untersuchungen von Geary, Hamson und Hoard (2000) für die erste und die zweite Klasse und von Jordan, Glutting und Ramineni (2010) für Ende erste und Ende dritte Klasse belegt werden. Allerdings zeigen sich für die Primarschule bei gewissen Aufgaben auch Deckeneffekte (Geary et al., 2000). Im Folgenden soll die Zählkompetenz als eine zentrale mathematische Vorläuferfertigkeit noch etwas ausführlicher thematisiert werden.

2.4.2 Zählkompetenz

Verschiedene Untersuchungen zeigen, dass rechenschwache Kinder grössere Schwierigkeiten haben, gewisse Zählfehler im Sinne der Zählprinzipien (Prinzip der stabilen Ordnung, Eindeutigkeitsprinzip, Kardinalprinzip, Abstraktionsprinzip, Prinzip der Irrelevanz der Anordnung) nach Gelman und Gallistel (1978) zu identifizieren, auch wenn sie die Zählprinzipien grundsätzlich verstehen (Geary, Bow-Thomas & Yao, 1992; Geary et al., 1999; Geary et al., 2007; Murphy et al., 2007). So zeigten rechenschwache Kinder schlechtere Leistungen bei der Anwendung des Prinzips der Irrelevanz der Anordnung (Geary et al., 1992) und bei der Beurteilung der Korrektheit eines Zählaktes, bei der das erste Objekt doppelt gezählt wurde (Geary et al., 2007). Letzteres könnte auch durch eine eingeschränkte Arbeitsgedächtniskapazität mitbedingt sein. Es wurde jedoch festgestellt, dass sich die Schwierigkeiten beim Identifizieren von Zählfehlern mit der Zeit abschwächen und es bereits in den ersten Jahren der Primarschule zu Deckeneffekten in entsprechenden Aufgaben kommt (Geary et al., 2004; Geary et al., 2007; Murphy et al., 2007). So schnitten rechenschwache Schülerinnen und Schüler des ersten Primarschuljahres unter Kontrolle der Intelligenz beim Identifizieren von Zählfehlern in der Untersuchung von Geary et al. (2004) zwar etwas schlechter ab als Schülerinnen und Schüler ohne mathematische Schwierigkeiten. Bei rechenschwachen Schülerinnen und Schülern des dritten und des fünften Schuljahres konnte dies jedoch nicht mehr bestätigt werden (Geary et al., 2004).

Die genannten Untersuchungen machen allerdings keine Aussagen dazu, wie gut rechenschwache Kinder Zählprozeduren selber durchführen können. Hierzu sind Studien interessant, die auch komplexere Zählkompetenzen wie das Zählen in Schritten grösser als eins oder das Rückwärtszählen untersuchen. So konnte Moser Opitz (2013) belegen, dass schwache Rechnerinnen und Rechner des fünften und des achten Schuljahres zwar beim Zählen in Einerschritten vorwärts und rückwärts gleich gut abschnitten wie Kinder ohne mathematische Schwierigkeiten, jedoch über alle Aufgaben zum Zählen hinweg signifikant schlechtere Leistungen zeigten als die Vergleichsgruppe. Die Schwierigkeiten konnten dabei auf das Zählen in Zweierschritten vorwärts und rückwärts und auf das Zählen in Hunderterschritten rückwärts zurückgeführt werden. Fehleranalysen weisen zudem darauf hin, dass das Zählen über den Zehnerübergang und den Hunderterübergang für rechenschwache Kinder besonders schwierig ist und sie dabei häufig Zahlen auslassen und die Schrittgrösse während des Zählens ändern (Moser Opitz, 2013).

2.4.3 Rechenstrategien: Faktenabruf und zählendes Rechnen

Studien zeigen, dass schwache Rechnerinnen und Rechner auch grosse Mühe beim Abruf von mathematischem Faktenwissen haben (z. B. Geary, 1993; Jordan & Hanich, 2003; Landerl et al., 2004). Zum basalen arithmetischen Faktenwissen gehört das Wissen um die Lösungen einfacher Additions- und Subtraktionsaufgaben im Zahlenraum bis 20 sowie einfacher Multiplikations- und Divisionsaufgaben (kleines Einmaleins und kleines Einsdurcheins; Ashcraft, 1982; Busch et al., 2013). Dieses Wissen eignen sich Schülerinnen und Schüler in den ersten Schuljahren an, indem sie Ergebnisse einfacher Grundoperationen automatisieren und im Langzeitgedächtnis abspeichern (Busch et al., 2013). Schwache Rechnerinnen und Rechner des dritten bis sechsten Primarschuljahres schnitten in einer Untersuchung von Andersson (2010) bei einfachen Subtraktions-, Additions- und Multiplikationsaufgaben, bei denen das Ergebnis aus dem Gedächtnis abgerufen werden sollte, signifikant schlechter ab und benötigten zudem signifikant mehr Zeit für den Faktenabruf. Zu vergleichbaren Ergebnissen für schwache Rechnerinnen und Rechner des dritten und des fünften Primarschuljahres kam auch eine Studie von Vanbist et al. (2014). Eine Studie von Busch et al. (2013) zeigt zudem, dass schwache Rechnerinnen und Rechner des dritten Primarschuljahres auch grössere Schwierigkeiten haben, die Korrektheit einfacher Additionsaufgaben im Zahlenraum bis 20 in kurzer Zeit zu beurteilen, und zwar dann, wenn diese Aufgaben einen Zehnerübergang beinhalten oder wenn dabei eine Parallelaufgabe bearbeitet werden muss, die das Arbeitsgedächtnis belastet. Andersson (2008a) konnte zudem nachweisen, dass Schwierigkeiten beim Faktenabruf dazu beitragen, dass mehrstellige Additionsaufgaben von schwachen Rechnerinnen und Rechnern schlechter gelöst werden. Interessanterweise konnten Geary et al. (2004) feststellen, dass rechenschwache Kinder bei komplexeren Additionsaufgaben (zweistelliger Summand + einstelliger Summand = ?) häufiger auf Faktenabruf zurückgreifen als die Vergleichsgruppe, dabei aber viele Fehler machen. Es wurde daher die Vermutung aufgestellt, dass diese Kinder geraten haben, statt auf sinnvollere Zähl- oder Dekompositionsstrategien (z.B. 5 + 6 = ?, 5 + 5 = 10, 10 + 1 = 11) zurückzugreifen.

Da rechenschwache Kinder grosse Mühe mit dem Faktenabruf haben, greifen sie häufiger und länger auf zählende Rechenstrategien zurück (z.B. Fingerzählen, verbales Zählen) als Kinder ohne mathematische Schwierigkeiten (Geary et al., 1992; Geary et al., 2004; Geary, Hoard & Bailey, 2012; Mazzocco et al., 2008; Moser Opitz, 2013; Ostad, 1997). Verfestigtes zählendes Rechnen wird dann angenommen, wenn Kinder auch nach Ende des ersten Schuljahres beim Lösen von Additions- und Subtraktionsaufgaben zu einem grossen Teil auf zählende Rechenstrategien zurückgreifen (Schipper, 2002). Ostad (1997) kommt aufgrund seiner Untersuchung zum Schluss, dass sich der Strategiengebrauch von Kindern ohne mathematische Schwierigkeiten über die Jahre hinweg grundsätzlich durch eine Abnahme zählenden Rechnens, eine Zunahme beim Faktenabruf und allgemein durch eine grössere Vielfalt der verwendeten Strategien auszeichnet. Rechenschwache Kinder hingegen rechnen über die Jahre hinweg primär zählend und eignen sich deutlich weniger neue Strategien an (Ostad, 1997). Auch Vanbist et al. (2014) folgern, dass rechenschwache Kinder einen wenig effizienten Strategiengebrauch aufweisen.

Für die erste Klasse konnten Studien belegen, dass zählende Rechenstrategien sowohl von Kindern mit mathematischen Schwierigkeiten als auch von Kindern ohne solche Schwierigkeiten ähnlich häufig verwendet werden, Erstere dabei jedoch mehr Fehler machen und etwas häufiger die Strategie „Fingerzählen“ verwenden (Geary et al., 1992; Geary et al., 2004). Zu Beginn der Schulzeit greifen somit auch Kinder ohne mathematische Schwierigkeiten auf zählende Rechenstrategien zurück. Dies kann als sinnvoll eingeschätzt werden, da die Möglichkeit des Faktenabrufs zu diesem Zeitpunkt noch begrenzt ist, weil erst wenig Faktenwissen aufgebaut werden konnte (Geary et al., 2004). Im zweiten Schuljahr verwenden rechenschwache Schülerinnen und Schüler hingegen signifikant häufiger die Strategie „Fingerzählen“ und „verbales Zählen“ und greifen seltener auf Faktenabruf und Dekompositionsstrategien zurück (Geary, Hoard

& Bailey, 2012). Bis zur vierten Klasse zeigt sich zwar für alle untersuchten Kindern ein Rückgang an Fingerzählstrategien und ein Zuwachs von Abrufstrategien, rechenschwache Kinder rechnen jedoch nach wie vor signifikant häufiger zählend und machen sowohl beim Abrufen als auch beim Fingerzählen signifikant mehr Fehler als Kinder ohne Rechenschwäche (Geary, Hoard & Bailey, 2012). Moser Opitz (2013) konnte auch für die fünfte Klasse belegen, dass schwache Rechnerinnen und Rechner häufiger zählende Rechenstrategien anwenden, dass die Häufigkeit bis zur achten Klasse jedoch etwas abnimmt, sodass vermutet wird, dass die Automatisierung von Zahlenfakten mit der Zeit auch bei rechenschwachen Schülerinnen und Schülern zunimmt. Häufigeres Zählen der Finger konnte für die fünfte Klasse auch von Geary et al. (2004) und zudem auch für die achte Klasse von Mazzocco et al. (2008) bestätigt werden. Zusammenfassend kann also gesagt werden, dass rechenschwache Kinder über einen eingeschränkten Strategiengebrauch verfügen, grosse Mühe mit dem Faktenabruf haben, viel länger auf aufwendige, zählende Rechenstrategien zurückgreifen und dabei häufiger Fehler machen als Kinder ohne mathematische Schwierigkeiten.

2.4.4 Operationsverständnis und Mathematisieren

Wie bereits im Kapitel 2.4.3 aufgezeigt werden konnte, haben rechenschwache Kinder grössere Mühe, Additions-, Subtraktions-, Multiplikations- und Divisionsaufgaben[7] korrekt zu lösen, als Kinder ohne mathematische Schwierigkeiten, d.h., sie weisen schlechtere prozedurale Fähigkeiten auf. Andersson (2008a) konnte zudem aufzeigen, dass Kinder mit einem Förderbedarf in Mathematik der dritten und der vierten Klasse zentrale Rechenprinzipien wie das Kommutativgesetz ($4 + 5 = 9$ ist das Gleiche wie $5 + 4 = 9$) oder das Prinzip der Umkehraufgaben ($4 + 5 = 9$, d.h. $9 – 5 = 4$) signifikant schlechter verstehen. Diese Ergebnisse wurden durch eine spätere Studie auch für die fünfte und die sechste Klasse bestätigt (Andersson, 2010). Auch Jordan et al. (2003) konnten für Ende der dritten Klasse belegen, dass rechenschwache Schülerinnen und Schüler die Rechenprinzipien signifikant schlechter verstehen und anwenden können. Es kann daher angenommen werden, dass Schülerinnen und Schüler mit einer Rechenschwäche über ein eingeschränktes Operationsverständnis verfügen.

Von grosser Wichtigkeit ist jedoch, dass rechenschwache Schülerinnen und Schüler auch weniger Verständnis dafür haben, welche mathematischen Handlungen hinter den einzelnen Grundoperationen stehen (Scherer & Moser Opitz, 2010). So konnte Moser Opitz (2013) aufzeigen, dass rechenschwache Schülerinnen und Schüler der vierten und der achten Klasse signifikant

[7] Multiplikations- und Divisionsaufgaben wurden in den genannten Studien eher selten erfasst. Belege für schlechtere Leistungen rechenschwacher Kinder konnten jedoch Andersson (2010) für die Multiplikation, Landerl, Bevan und Butterworth (2004) für die Division und Moser Opitz (2013) für beide Operationen belegen.

mehr Mühe haben, eine Rechnung mit Material oder als Zeichnung darzustellen oder eine passende Rechengeschichte zu erfinden. Diese Schwierigkeiten zeigten sich am deutlichsten bei Divisions-, Ergänzungs- und Multiplikationsaufgaben und in geringerer Ausprägung bei Subtraktionsaufgaben. Damit wird auf die Wichtigkeit von Mathematisierungskompetenzen verwiesen, die für das Lösen von Sach- und Textaufgaben benötigt werden, bei denen numerische und verbale Informationen zu verstehen, in eine innere Repräsentation zu übersetzen und darauf basierend Lösungen zu generieren sind (Moser Opitz, 2013). In den bereits genannten Untersuchungen von Andersson (2008a, 2010) hatten rechenschwache Kinder der dritten bis sechsten Klasse sowohl beim Lösen von Textaufgaben, die nur einen Rechenschritt verlangten, als auch von solchen, die mehrere Rechenschritte umfassten, signifikant schlechtere Leistungen im Vergleich zu Kindern ohne Rechenschwäche. Signifikant schlechtere Leistungen zeigten sich ebenfalls in der Studie von Jordan et al. (2003) und Jordan und Hanich (2003) für die zweite und die dritte Klasse. Auch für ältere Schülerinnen und Schüler mit einer Rechenschwäche der fünften und der achten Klasse konnten solche Befunde bestätigt werden (Moser Opitz, 2013). Dabei stellte sich heraus, dass in beiden Klassenstufen rechenschwache Schülerinnen und Schüler Austauschaufgaben bei der Addition und der Subtraktion recht gut lösen konnten (z.B. „P. hat 42 CDs. Er gibt 5 CDs seiner Freundin A. Wie viele CDs hat er noch?“; Moser Opitz, 2013, S. 203). Signifikante Leistungsunterschiede ergaben sich jedoch bei Vergleichsaufgaben sowohl für die Addition als auch für die Subtraktion, bei denen die Relation zwischen zwei Zahlen bestimmt werden muss, was als anspruchsvoller eingeschätzt werden kann (z.B. „T. hat eine CD-Sammlung. Er gibt L. 6 CDs. Nun bleiben ihm 37 CDs. Wie viele CDs hatte T. am Anfang?“; Moser Opitz, 2013, S. 203). Ebenfalls grosse Mühe hatten die rechenschwachen Kinder bei einer Divisionsaufgabe mit dem Sachkontext „Geld“, bei der ausgehend vom Preis für drei Nussgipfel der Preis für einen Nussgipfel bestimmt werden musste, wobei es vielen rechenschwachen Schülerinnen und Schülern nicht einmal gelang, die passende Divisionsrechnung zu notieren (Moser Opitz, 2013). Es kann somit davon ausgegangen werden, dass rechenschwache Schülerinnen und Schüler grössere Schwierigkeiten beim Mathematisieren haben und über ein eingeschränktes Operationsverständnis verfügen, was sich negativ auf prozedurale Fähigkeiten wie das Lösen von Sach- und Textaufgaben auswirkt.

2.4.5 Dezimales Stellenwertsystem

Einige Studien zeigen, dass rechenschwache Schülerinnen und Schüler auch Schwierigkeiten in Zusammenhang mit dem dezimalen Stellenwertsystem haben (Hanich, Jordan, Kaplan & Dick, 2001; Jordan et al., 2003; Jordan & Hanich, 2003; Moser Opitz, 2013). So konnten Hanich et al. (2001) belegen, dass rechenschwache Kinder der zweiten Klasse signifikant niedrigere Leistungen in Aufgaben zeigen, die das Verständnis für das dezimale Stellenwertsystem prüfen (z.B. Stellenwertschreibweise, Zahlen benennen). Auch Jordan und Hanich (2003) konnten sowohl für die zweite als auch die dritte Klasse bestätigen, dass rechenschwache Kinder signifikant schlechtere Leistungen in diesem Bereich aufweisen, und zwar auch dann, wenn der IQ kontrolliert wird. Auch für die dritte Klasse konnte dies in einer weiteren Untersuchung belegt werden, allerdings war der Unterschied zwischen Schülerinnen und Schülern mit und ohne Rechenschwäche hier nicht mehr signifikant, wenn zusätzlich zum IQ auch noch das Geschlecht, die ethnische Herkunft und das familiäre Einkommen kontrolliert wurden (Jordan et al., 2003). Untersuchungen älterer Schülerinnen und Schüler zeigen jedoch, dass auch in der fünften und der achten Klasse eine Rechenschwäche mit niedrigeren Leistungen im Bereich des dezimalen Stellenwertsystems einhergeht, wobei dies auf mangelndes konzeptuelles Verständnis zurückgeführt wird (Moser Opitz, 2013). Mazzocco et al. (2008) konnten entsprechende Schwierigkeiten in der achten Klasse ebenfalls bestätigen, allerdings nur für rechenschwache Schülerinnen und Schüler mit sehr tiefen Mathematikleistungen (< 10. Prozentrang). Zudem konnte eine Langzeitstudie aufzeigen, dass das Verständnis des Stellenwertsystems in der ersten Klasse einen signifikanten Einfluss auf die Leistung in Additionsaufgaben in der dritten Klasse hat und somit für die weitere mathematische Entwicklung von grosser Wichtigkeit ist, unabhängig davon, ob ein Kind eine Rechenschwäche aufweist oder nicht (Moeller et al., 2011). Im Widerspruch zu den bisher genannten Studien stehen hingegen die Ergebnisse von Andersson (2008a, 2010). In beiden Untersuchungen haben sich für rechenschwache Schülerinnen und Schüler der dritten und der vierten Klasse keine schlechteren Leistungen in Aufgaben zum dezimalen Stellenwertsystem gezeigt. Allerdings sind die dafür eingesetzten Aufgaben als eher einfach einzuschätzen, sodass angenommen werden kann, dass diese Aufgaben nicht ein tiefergehendes Verständnis des Dezimalsystems prüfen, sondern eher oberflächliche Kenntnisse erfassen (vgl. Freesemann, 2014). Es kann daher trotz teilweise widersprüchlicher Ergebnisse davon ausgegangen werden, dass rechenschwache Schülerinnen und Schüler auch ein eingeschränktes Verständnis des dezimalen Stellenwertsystems haben, was zu geringeren Leistungen in entsprechenden Aufgaben führt.

3. Fördermassnahmen bei Rechenschwäche

Im letzten Kapitel wurde aufgezeigt, welche Inhaltsbereiche für eine mathematische Förderung zu berücksichtigen sind, da rechenschwache Schülerinnen und Schüler in diesen grosse Schwierigkeiten aufweisen. Inzwischen gibt es auch vermehrt Interventionsstudien zur Wirksamkeit von Förderansätzen bei Rechenschwäche, die ebenfalls wichtige Hinweise für die Gestaltung einer Mathematikförderung bieten. Im Folgenden werden daher zunächst Metastudien und verstehensorientierte Interventionsstudien für die Förderung bei Rechenschwäche vorgestellt. Im Anschluss daran werden didaktische sowie mathematikdidaktische Erkenntnisse zur Förderung rechenschwacher Lernender präsentiert. Dabei spielen didaktische Überlegungen zur Strukturierung des Unterrichts mit Blick auf rechenschwache Schülerinnen und Schüler eine wichtige Rolle, ebenso stellt sich aber auch die Frage nach der Rolle gemeinsamer Lernsituationen, wenn eine Förderung in einem integrativen Setting umgesetzt werden soll. Die Mathematikdidaktik bietet wiederum zahlreiche Hinweise, was bei der Gestaltung eines guten Mathematikunterrichts unter Berücksichtigung rechenschwacher Lernender sowohl inhaltsübergreifend als auch bei der Erarbeitung spezifischer Inhalte wichtig ist. Das Kapitel schliesst mit Überlegungen zur Bedeutsamkeit von professionellem Wissen bzw. von professionellen Kompetenzen von Lehrpersonen, die für diese Arbeit ebenfalls zu beachten sind.

3.1 Evaluierte Förderansätze

Es gibt inzwischen einige empirische Untersuchungen sowie auch Metastudien zur Förderung rechenschwacher Schülerinnen und Schüler. Ausser bei Metastudien werden im Folgenden nur Studien präsentiert, die eine verstehensorientierte Förderung arithmetischen Basiswissens beinhalten. D.h. Studien, die primär auf die Automatisierung von Faktenwissen abzielen (z.B. Codding, Burns & Lukito, 2011; Fuchs et al., 2008), oder solche, die sich ausschliesslich mit der Förderung von Problemlösekompetenzen auseinandersetzen (z.B. Fuchs & Fuchs, 2005; Montague, Krawec, Enders & Dietz, 2014; Swanson, Orosco & Lussier, 2014), werden hier nicht beschrieben. Zudem werden auch computerbasierte Förderprogramme im Folgenden nicht berücksichtigt (z. B. Kucian, Grond et al., 2011; Rauscher et al., 2016).

3.1.1 Metastudien

In einer sehr umfassenden Metastudie von Gersten et al. (2009) wurden 41 mathematische Interventionsstudien mit randomisiert kontrollierten sowie mit hochwertigen quasi-experimentellen Designs bezüglich ihrer Effektstärken und den verwendeten didaktischen Ansätzen verglichen. Dabei wurden nur Studien berücksichtigt, die Kinder mit einer „mathematical learning disability“ (d.h. Kinder mit einer Rechenstörung, die das doppelte Diskrepanzkriterium erfüllen) untersucht und die ein klar definiertes didaktisches Konzept für die mathematische Intervention beinhaltet haben. Es zeigte sich, dass Ansätze der expliziten Instruktion (durchschnittliche Effektstärke Hedges g = 1.22; niedrigster Wert: 0.08; höchster Wert: 2.15; N = 11), die Verwendung von Heuristiken (Hedges g = 1.56; niedrigster Wert: 0.54; höchster Wert: 2.45; N = 4) und das Anregen der Schülerinnen und Schüler zum Verbalisieren (durchschnittliche Effektstärke Hedges g = 1.04; niedrigster Wert: 0.07; höchster Wert: 2.01; N = 8) besonders wirksam waren. Ebenfalls als effektiv erwiesen sich didaktische Ansätze, in denen Aufgaben speziell ausgewählt oder sequenziert (durchschnittliche Effektstärke Hedges g = 0.82; niedrigster Wert: 0.12; höchster Wert: 2.15; N = 9) und in denen Lehrpersonen und/oder Lernenden Feedbacks zu den Leistungen der Schülerinnen und Schüler gegeben wurden (Hedges g zwischen 0.21 und 0.34), sowie Ansätze, in denen gezielt mit Visualisierungen gearbeitet (durchschnittliche Effektstärke Hedges g = 0.47; niedrigster Wert: –0.29; höchster Wert: 1.39; N = 12) oder in denen ein Tutorium durch ältere Tutorinnen oder Tutoren gegeben wurde (durchschnittliche Effektstärke Hedges g = 1.02; niedrigster Wert: 0.75; höchster Wert: 1.15; N = 2). In dieser Metastudie wurde zudem auch der relative Anteil der einzelnen Ansätze untersucht, da in einigen Studien verschiedene Ansätze in Kombination zur Anwendung kamen. Dabei zeigte sich, dass das Verwenden von Heuristiken und die explizite Instruktion besonders wirksame Strategien für die Förderung von Schülerinnen und Schülern mit einer Rechenstörung sind.

Zu vergleichbaren Ergebnissen kam eine ältere, internationale Metastudie, in der auch Schülerinnen und Schüler mit Rechenschwäche berücksichtigt wurden, die das doppelte Diskrepanzkriterium nicht erfüllen (Kroesbergen & van Luit, 2003). Auch hier zeigte sich, dass das Verwenden von Heuristiken (gewichtete Effektstärke d = 1.45) und Ansätze der direkten Instruktion (gewichtete Effektstärke d = 0.91) für rechenschwache Lernende die effektivsten Mittel für eine mathematische Förderung waren.

Weitere interessante Hinweise bietet eine Metaanalyse acht deutschsprachiger Studien, bei denen die Förderung von rechenschwachen Schülerinnen und Schülern untersucht wurde (Ise et al., 2012). Die durchschnittliche Effektstärke der berücksichtigten Studien betrug 0.50, und es zeigte sich, dass sowohl curricular ausgerichtete Förderansätze (durchschnittliche Effektstärke Hedges g = 0.50) als auch entwicklungspsychologisch (durchschnittliche Effektstärke Hedges g = 0.43) bzw. neuropsychologisch (Effektstärke Hedges g = 1.08, nur eine Studie) orientierte Förderansätze effektiv waren (Ise et al., 2012). Zudem ergaben die Analysen, dass eine Einzelförderung (durchschnittliche Effektstärke Hedges g = 1.39) deutlich effektiver war als eine Gruppenförderung (durchschnittliche Effektstärke Hedges g = 0.28) bzw. eine klassenweise

Förderung (durchschnittliche Effektstärke Hedges $g = 0.21$). Allerdings konnten bei der Einzelförderung und der klassenweisen Förderung jeweils nur zwei Untersuchungen für die Metastudie berücksichtigt werden. Zudem zeigte sich, dass rechenschwache Schülerinnen und Schüler von einer Förderung, die länger als zehn Wochen dauerte, stärker profitieren konnten (durchschnittliche Effektstärke Hedges $g = 1.39$) als von einer Förderung, die kürzer als zehn Wochen war (durchschnittliche Effektstärke Hedges $g = 0.33$).

Zusammenfassend kann festgehalten werden, dass die genannten Metastudien die Wirksamkeit mathematischer Förderung für rechenschwache Schülerinnen und Schüler belegen können, dass je nach Charakteristika der Förderung (z.B. Förderansatz, Setting, Dauer) jedoch unterschiedlich hohe Effektstärken festgestellt werden. Mit Blick auf die verwendeten Förderansätze haben sich insbesondere stark strukturierte Instruktionsansätze zumindest kurzfristig als sehr wirksam erwiesen. Wie solche Ansätze zur Strukturierung des Mathematikunterrichts konkret aussehen und was an diesen auch kritisch zu betrachten ist, wird in der vorliegenden Arbeit in Zusammenhang mit didaktischen Überlegungen im Kapitel 3.2 noch genauer erläutert. Auch das Anregen der Schülerinnen und Schüler zum Verbalisieren, eine produktiv gestaltete Übungspraxis sowie der Einsatz von Visualisierungen, die sich ebenfalls als wirksam für die Förderung bei Rechenschwäche erwiesen haben, werden in dieser Arbeit noch genauer ausgeführt (vgl. Kap. 3.3). Im Folgenden wird zunächst jedoch auf neuere Interventionsstudien eingegangen, die auf eine verstehensorientierte Förderung arithmetischen Basiswissens für rechenschwache Schülerinnen und Schüler fokussieren und ein quasi-experimentelles Design aufweisen.

3.1.2 Interventionsstudien zur Förderung von Vorläuferfertigkeiten

Es gibt einige Studien, die die Wirksamkeit einer Förderung von Vorläuferfertigkeiten bei rechenschwachen Kindern im Sinne des Entwicklungsmodells der Zahl-Grössen-Verknüpfung gemäss Krajewski (2003, 2005, 2007) untersucht haben (für eine Übersicht vgl. Ennemoser et al., 2015). Im Folgenden werden nur Studien erwähnt, in denen die Förderung in der Primarschule durchgeführt wurde.

Eine erste Studie zur Wirksamkeit eines Trainings von Vorläuferfertigkeiten im Sinne des Entwicklungsmodells der Zahl-Grössen-Verknüpfung von Krajewski (2003, 2005, 2007) für rechenschwache Kinder wurde von Ennemoser und Krajewski (2007) mit einer kleinen Stichprobe ($N = 30$) in der ersten Klasse in Deutschland durchgeführt. Für die Untersuchung wurden nur Schülerinnen und Schüler mit einer Rechenschwäche berücksichtigt (PR < 25 im DEMAT 1+; Krajewski, Küspert & Schneider, 2002). Die Förderung für die Interventionsgruppe war kompetenz- bzw. entwicklungsorientiert. Inhaltlich fokussierte sie auf das Teile-Ganzes-

Verständnis (Ebene III gemäss Entwicklungsmodell), es wurden Arbeitsgedächtnis-entlastende sowie strukturorientierte Veranschaulichungen und Arbeitsmittel eingesetzt und auf eine sprachliche Begleitung mathematischer Handlungen geachtet. Die Förderung wurde durch geschulte Projektmitarbeitende in Kleingruppen mit fünf Kindern durchgeführt und umfasste insgesamt sechs Schulstunden verteilt auf drei Wochen. Es zeigte sich, dass das eingesetzte Training der Vorläuferfertigkeiten wirksam war, bei einer um Vortestunterschiede korrigierten Effektstärke von $d = 0.58$. Gemäss Cohen (1988) handelt es sich hierbei um eine mittlere Effektstärke. Die positiven Ergebnisse für rechenschwache Kinder wurden als Transfereffekt interpretiert, da im verwendeten Mathematiktest auch Rechenaufgaben geprüft wurden, diese hingegen nicht Teil der Förderung der Vorläuferfertigkeiten waren. Längerfristige Effekte wurden in dieser Studie allerdings nicht untersucht.

In einer weiteren Studie wurde an einer grösseren Stichprobe mit rechenschwachen Kindern ($N = 119$) ebenfalls ein Training von Vorläuferfertigkeiten gemäss dem Entwicklungsmodell von Krajewski (2003, 2005, 2007) untersucht (Sinner, 2011)[8]. Für das Training wurde das Förderprogramm „Mengen, zählen, Zahlen“ (MZZ; Krajewski, Nieding & Schneider, 2007) für den Einsatz in der ersten Klasse adaptiert. Als Cut-off-Wert für das Vorliegen einer Rechenschwäche wurde ein Prozentrang von unter 20 im MBK 1+ (Ennemoser et al., 2017) gewählt, der frühe mathematisch Basiskompetenzen erfasst. Für die Überprüfung von Transfereffekten auf mathematische Leistungen im engeren Sinne wurde zusätzlich beim Post-Test und beim Follow-up 1 der DEMAT 1+ (Krajewski et al., 2002) sowie beim Follow-up 2 Ende zweiter Klasse der HRT 1–4 (Haffner, Baro, Parzer & Resch, 2005) eingesetzt. Die Förderung fand im dritten Quartal der ersten Klasse in Kleingruppen von zwei bis sieben Kindern statt und umfasste zwölf Sitzungen. Die Förderung fokussierte dabei auf die Ebene II und III des Entwicklungsmodells von Krajewski (2003, 2005, 2007). Die um Vortestunterschiede bereinigte Effektstärke des Trainings lag für mathematische Vorläuferfertigkeiten, erfasst mit dem Gesamttest des MBK 1+ (Ennemoser et al., 2017), beim Post-Test bei $d = 1.34$ und beim Follow-up 1 bei $d = 1.24$. Transfereffekte auf den standardisierten Mathematiktest konnten für den Post-Test jedoch nicht gefunden werden. Erst beim Follow-up 1 zeigte sich, dass die Schülerinnen und Schüler der Interventionsgruppen auch im standardisierten Mathematiktest signifikant von der Förderung profitiert haben (um Vortestunterschiede bereinigte Effektstärke: $d = 0.77$). Zusätzlich konnte auch für den Follow-up 2 ein signifikanter Transfereffekt festgestellt werden, jedoch bei einer kleineren, ebenfalls um Vortestunterschiede bereinigten Effektstärke von $d = 0.37$. Da Transfereffekte erst beim Follow-up 1 belegt werden konnten, wird vermutet, dass die

[8] Berichtet werden hier nur das Design und die Ergebnisse der Hauptstudie von Sinner (2011).

rechenschwachen Kinder zunächst Lücken in den Vorläuferfertigkeiten schliessen mussten, bevor sich das Trainingsprogramm auf die mathematischen Leistungen im engeren Sinne auswirken konnte.

In einer weiteren Untersuchung von Ennemoser et al. (2015) wurde geprüft, ob eine Förderung von Vorläuferfertigkeiten auf Basis des Entwicklungsmodells von Krajewski (2003, 2005, 2007) auch bei sogenannten Risikokindern in der ersten Klasse effektiv ist (N = 64). Mit Risikokindern sind hier Schülerinnen und Schüler gemeint, die ein erhöhtes Risiko aufweisen, eine Rechenschwäche zu entwickeln (PR < 25 im MBK 1+; Ennemoser et al., 2017). Auch hier wurden Einheiten des Förderprogramms MZZ auf Ebene II und III eingesestzt (Krajewski et al., 2007). Mathematische Rechenbegriffe (plus, minus, gleich) und Rechenaufgaben auf Symbolebene waren hier hingegen nicht Teil der Förderung. Die Förderung wurde von Lehramtsstudierenden in Kleingruppen von drei bis sieben Kindern durchgeführt und umfasste zehn Sitzungen. Um Transfereffekte zu prüfen, wurden zusätzlich auch Rechenfertigkeiten unter Zeitdruck erfasst. Es zeigte sich beim Post-Test mit Blick auf die Vorläuferfertigkeiten kein Haupteffekt für die Gruppenzugehörigkeit, allerdings war der Lernzuwachs in der Interventionsgruppe signifikant grösser als in der Kontrollgruppe (um Vortestunterschiede bereinigte Effektstärke von d = 0.64). Beim Follow-up waren die Ergebnisse vergleichbar (um Vortestunterschiede bereinigte Effektstärke von d = 0.69). Mit Bezug auf die Rechenfertigkeiten ergab sich beim Post-Test kein Effekt der Förderung. Beim Follow-up konnte jedoch ein Haupteffekt für die Versuchsbedingung gefunden werden (um Vortestunterschiede bereinigte Effektstärke von d = 0.52). Auch in dieser Untersuchung konnten Transfereffekte somit erst für den Follow-up festgestellt werden.

3.1.3 Interventionsstudien zur Förderung arithmetischer Kompetenzen in der Unterstufe

Aus den USA liegt eine Studie vor, in der eine verstehensorientierte mathematische Förderung für rechenschwache Kinder in der ersten (N = 126, $n_{\text{Intervention}}$ = 26) und der zweiten Klasse (N = 140, $n_{\text{Intervention}}$ = 25) untersucht wurde, die auf frühe arithmetische Kompetenzen fokussiert (Bryant et al., 2008). In dieser wurden der Aufbau eines tragfähigen Zahlverständnisses (u.a. Zählen, Grössenvorstellung, Teile-Ganzes-Verständnis) und das Verständnis des dezimalen Stellenwertsystems gefördert, und es wurden Strategien für einfache Additions- und Subtraktionsaufgaben erarbeitet. Dabei wurde ein expliziter und systematischer Unterrichtsstil verwendet, und Lösungsstrategien wurden vorgezeigt. Mithilfe von geeigneten Arbeitsmitteln und Veranschaulichungen wurden Verknüpfungen zwischen verschiedenen Repräsentationsebenen geübt. Die Förderung fand in Kleingruppen von zwei bis vier Kindern in Form von 15-

minütigen Lektionen über einen Zeitraum von 18 Wochen statt (ca. 62 bis 64 Lektionen) und wurde von speziell geschulten Tutoren mit pädagogischem Hintergrund durchgeführt. Als Cut-off-Wert für das Vorliegen einer Rechenschwäche wurde ein Prozentrang von unter 25 im verwendeten Mathematiktest gewählt (TEMI-PM; University of Texas System & Texas Education Agency, 2007), der frühe arithmetische Kompetenzen unter Zeitdruck überprüft. Die Ergebnisse zeigen für das zweite Primarschuljahr, nicht jedoch für das erste Primarschuljahr einen signifikanten, positiven Effekt der Förderung. Es wird vermutet, dass der nicht vorhandene Effekt bei den jüngeren Kindern durch eine zu geringe Intensität und Dauer der Förderlektionen bedingt sein könnte. Hinweise zur Bestätigung dieser Annahme konnten in einer neueren Untersuchung gefunden werden (Bryant et al., 2011). In dieser Studie wurde derselbe Förderansatz verwendet, die Förderung dauerte jedoch länger und beinhaltete 25-minütige Lektionen während 19 Wochen (76 Lektionen). Es zeigte sich, dass unter diesen Bedingungen auch rechenschwache Schülerinnen und Schüler der ersten Klasse signifikant von der Förderung profitieren konnten (TEMI-PM). Transfereffekte auf weiterführende Inhalte (u.a. Problemlösen) wurden hier ebenfalls untersucht, allerdings konnten diese nicht bestätigt werden.

In einer weiteren Studie (N = 64) wurde ein an neuropsychologischen Modellen orientiertes Förderprogramm für rechenschwache Schülerinnen und Schüler in der zweiten und der dritten Klasse untersucht, das ebenfalls verstehensorientiert war (Wißmann et al., 2013). Die Förderung wurde hier in jahrgangsgemischten Kleingruppen von zwei bis sechs Kindern durchgeführt und umfasste ein ganzes Schuljahr, wobei pro Woche eine Doppellektion zur Verfügung stand. Es wurde ein systematischer Aufbau basisnumerischer und konzeptueller Kompetenzen verfolgt und prozedurales und arithmetisches Faktenwissen gefördert. Das Förderprogramm (vgl. Kaufmann, Handl & Thöny, 2003) orientiert sich am Triple-Code-Modell der Zahlenverarbeitung (Dehaene & Cohen, 1995) und beinhaltet Module zur Mengenerfassung, zum Zählen, zu den vier Grundoperationen, zum Teile-Ganzes-Verständnis sowie zum dezimalen Stellenwertsystem. Die Module starten jeweils auf einer anschaulichen, handlungsorientierten Ebene unter Verwendung von geeigneten Materialien, und es erfolgt ein schrittweiser Aufbau zu symbolischen Darstellungen bis hin zur Automatisierung. Die Förderung fand durch Mitarbeitende eines psychologischen Instituts statt, wobei stark standardisierte Skripte für die Umsetzung der Förderung eingesetzt wurden. Als Kriterium für eine Rechenschwäche wurde verlangt, dass in zwei unterschiedlichen Mathematiktests niedrige Leistungen erzielt wurden (PR $\leq$ 20 im HRT 1-4; Haffner et al., 2005, und im TEDI-MATH; Kaufmann et al., 2008). Die Ergebnisse ergaben, dass rechenschwache Schülerinnen und Schüler, die an der mathematischen Förderung teilgenommen haben, signifikant bessere Leistungen erreichten. Dabei zeigte sich, dass Schülerinnen und Schüler mit einer isolierten Rechenschwäche stärker von der Förderung profitieren konnten als Schülerinnen und Schüler, die zusätzlich zu einer Rechenschwäche auch eine Leseschwäche hatten.

3.1.4 Interventionsstudien zur Förderung arithmetischer Kompetenzen in der Mittelstufe

Auch für die in der Schweiz als Mittelstufe bezeichneten Schuljahre konnten zwei bedeutsame Studien gefunden werden, in denen eine verstehensorientierte Förderung umgesetzt wurde. In einer dieser Interventionsstudien ($N = 123$) wurde an einer deutschen Stichprobe untersucht, ob eine Förderung konzeptuellen Verständnisses im Bereich des arithmetischen Basiswissens bei Kindern mit Rechenschwäche effektiv ist (Freesemann, 2014; Moser Opitz et al., 2017). Inhaltlich fokussierte die Förderung auf das Verständnis des Dezimalsystems und das Verständnis der Grundoperationen, zusätzlich wurden aber auch prozedurale Kompetenzen und der Aufbau von Faktenwissen gefördert. Zudem wurden geeignete Veranschaulichungen und Hilfsmittel eingesetzt, auf eine individuelle Lernbegleitung geachtet, Abstraktionsprozesse gefördert, die Versprachlichung von Strategien, Erklärungen und Prozeduren geübt sowie auf eine klare Strukturierung der einzelnen Fördereinheiten berücksichtigt. An der Untersuchung haben sowohl Schülerinnen und Schüler der fünften Klasse aus Regelschulen teilgenommen als auch Schülerinnen und Schüler der siebten Klasse, die eine Schule für den Förderschwerpunkt Lernen besucht haben und bezüglich Curricula als vergleichbar angesehen wurden. Als Kriterium für eine Rechenschwäche mussten die Schülerinnen und Schüler Mathematikleistungen aufweisen, die um mindestens eine Standardabweichung niedriger waren als der Mittelwert der Normstichprobe (BASIS-MATH-G 4^{+}–5; Moser Opitz, Freesemann, Grob & Prediger, 2016). Es wurden zwei unterschiedliche Interventionssettings untersucht, wobei in beiden Fällen angehende Lehrpersonen die Förderung während vierzehn Wochen umgesetzt haben. In der einen Interventionsgruppe wurde in Kleingruppen von drei bis fünf Schülerinnen und Schülern während 90 Minuten pro Woche das Förderprogramm umgesetzt, während der restlichen Lektionen besuchten die Schülerinnen und Schüler den regulären Mathematikunterricht der Klasse. Die andere Interventionsgruppe wurde während 45 Minuten pro Woche in Kleingruppen von drei bis fünf Kindern gefördert und konnte zusätzlich während weiterer 90 Minuten im regulären Mathematikunterricht individuell an den Inhalten der Intervention arbeiten, sodass von einer teilweise klassenintegrierten Förderung gesprochen wird. Die Interventionsgruppen haben mit den gleichen Materialien und Aufgaben gearbeitet, und es wurden vergleichbare Aktivitäten durchgeführt, die Lektionen waren jedoch nicht identisch aufgebaut (vgl. Freesemann, 2014, S. 159ff). Sowohl für den Post-Test als auch für den Follow-up konnte für die teilweise klassenintegrierte Förderung festgestellt werden, dass die rechenschwachen Schülerinnen und Schüler signifikant von der Förderung profitiert haben und sich auch Transfereffekte bei nicht geförderten Inhalten ergeben haben. Die Effektstärken liegen für die geförderten Inhalte bei 0.75 *SD* (Post-Test) bzw. bei 0.54 *SD* (Follow-up), die Effektstärken für den Gesamttest bei 0.32 *SD* (Post-Test) bzw. 0.48 *SD* (Follow-up; vgl. Moser Opitz et al., 2017). Entgegen den Erwartungen hat die Interventionsgruppe, in der die gesamte Förderung in Kleingruppen stattgefunden hat, jedoch bei beiden Messzeitpunkten nur im Bereich der geförderten Inhalte signifikant von der Förderung profitiert. Die Effektstärken lagen hier bei 0.50 *SD* (Post-Test) bzw. 0.46 *SD* (Follow-up). Bezüglich der Leistungen der beiden Interventionsgruppen konnten hingegen keine signifikanten Unterschiede gefunden werden.

In einer weiteren Studie ($N = 53$) wurde untersucht, ob sich ein anspruchsvoller Mathematikunterricht in der sechsten Klasse (USA), der auf konzeptuelles Verständnis, Problemlösen und Sprechen über Mathematik fokussiert, wirksam ist (Woodward & Brown, 2006). Im Vordergrund des Forschungsinteresses stand die Frage, ob eine solche Förderung auch im natürlichen Kontext einer Schulklasse erfolgreich umgesetzt werden kann. Dafür wurden Schulklassen ausgewählt, in denen ausschliesslich Schülerinnen und Schüler mit „learning disabilities“[9] unterrichtet wurden, die meistens im Bereich des Lesens individuelle Lernziele aufwiesen und bezüglich ihrer Mathematikleistung von den Lehrpersonen als Risikokinder eingeschätzt wurden, die jedoch in diesem Fach keine individuellen Lernziele zugewiesen bekommen haben. Für die Untersuchung wurde ein Förderprogramm entworfen, das aus einer 55-minütigen, täglichen Lektion bestand, in der zentrale curriculare Inhalte der sechsten Klasse behandelt wurden. Neue Konzepte wurden dabei im Klassenverband erarbeitet und es wurden sorgfältig gestaltete Übungsgelegenheiten angeboten, die neben Rechenfertigkeiten und Problemlösekompetenzen insbesondere auch konzeptuelle Einsicht fördern sollten, gefolgt von Reflexionsrunden im Klassenverband. Es wurden jeweils geeignete Veranschaulichungen und Hilfsmittel eingesetzt und anspruchsvolle, aber zu bewältigende Aufgaben verwendet. An der Untersuchung haben sechs erfahrene Lehrpersonen aus zwei vergleichbaren Schulhäusern mit ihren Schulklassen teilgenommen. Das eine Schulhaus diente als Interventionsgruppe ($n = 25$), die andere Schule als Kontrollgruppe ($n = 28$), wobei insgesamt nur zwei Lehrpersonen mit ihren Klassen die Förderung nach vorgegebenem Plan umgesetzt haben. Es zeigte sich beim Post-Test, dass die Schülerinnen und Schüler der Interventionsgruppe bei einer sehr hohen Effektstärke von $d = 1.23$ einen signifikant höheren Lernzuwachs aufwiesen als die Schülerinnen und Schüler der

[9] Hierfür mussten bundesstaatliche Kriterien für eine „learning disability“ erfüllt sein, die jedoch im Artikel nicht ausgeführt werden. Keines der Kinder hatte eine unterdurchschnittliche Intelligenz, und alle erfüllten ein nicht genauer beschriebenes Diskrepanzkriterium (Woodward & Brown, 2006, S. 153).

Kontrollgruppe (CTB Terra Nova; McGraw-Hill, 2002). Auch mit Bezug auf einen selbst entwickelten Test, der das Verständnis zentraler mathematischer Konzepte untersuchen soll, zeigten die Schülerinnen und Schüler signifikant höhere Leistungen bei einer ebenfalls als sehr hoch einzuschätzenden Effektstärke von $d = 1.95$.

Die genannten Interventionsstudien belegen somit, dass ein verstehensorientiertes Training wichtiger mathematischer Vorläuferfertigkeiten bzw. eine verstehensorientierte Förderung zentraler arithmetischer Kompetenzen bei rechenschwachen Schülerinnen und Schülern auf verschiedenen Klassenstufen wirksam sein kann. Alle Untersuchungen haben dabei Instruktionsmassnahmen eingesetzt, die sich im Rahmen der zitierten Metastudien als wirksam erwiesen haben (z.B. Strukturierung, Veranschaulichungen, Versprachlichung mathematischer Handlungen, sinnvoller Einsatz von Übungsgelegenheiten; vgl. Kap. 3.1.1). Abschliessend muss erwähnt werden, dass nur für die Mittelstufe Untersuchungen vorliegen, die eine solche Förderung (zumindest teilweise) in einem Klassensetting umgesetzt haben, wobei jedoch in beiden Studien nur rechenschwache Schülerinnen und Schüler an der Förderung teilgenommen haben und die Erkenntnisse somit nur sehr eingeschränkt auf ein reguläres Klassensetting übertragen werden können (Moser Opitz et al., 2017; Woodward & Brown, 2006).

3.2 Didaktische Überlegungen zur Förderung rechenschwacher Schülerinnen und Schüler

Wie im vorangehenden Kapitel aufgezeigt wurde, gelten strukturierte Förderprogramme gemäss Metaanalysen zu den wirksamsten Förderansätzen bei Rechenschwäche. Allerdings werden insbesondere Ansätze der direkten und der expliziten Instruktion auch kritisch gesehen, da sie stark auf Auswendiglernen fokussieren und primär auf Rechenfertigkeiten abzielen (Montague, 2011; Moser Opitz et al., 2017). Gefordert wird daher, dass eine mathematische Förderung vermehrt auf Einsicht und Verständnis abzielen müsse, dabei jedoch ausreichend Strukturierung aufweisen soll, um für rechenschwache Schülerinnen und Schüler geeignet zu sein (Montague, 2011; Scherer & Moser Opitz, 2010). Daher wird auf den nächsten Seiten genauer aufgezeigt, welche Vorgehensweisen in der direkten und der expliziten Instruktion verwendet werden und was daran als kritisch erachtet wird. Darauf aufbauend wird diskutiert, welche Massnahmen sich für eine Strukturierung einer verstehensorientierten Mathematikförderung eignen, wobei insbesondere auf eine adaptive Lernbegleitung eingegangen wird. In einem zweiten Schritt wird zudem der Frage nach der Bedeutsamkeit gemeinsamer Lernsituationen für die Förderung rechenschwacher Schülerinnen und Schüler im regulären Mathematikunterricht nachgegangen, da ein Mangel an gemeinsamen Lernsituationen aus verschiedenen Gründen als problematisch angesehen wird. Sowohl mit Blick auf Strukturierungsmassnahmen als auch mit Blick auf gemeinsame Lernsituationen geht es im Folgenden also um didaktische Überlegungen, die für die Gestaltung einer verstehensorientierten Förderung rechenschwacher Schülerinnen und Schüler im regulären Mathematikunterricht wichtige Hinweise bieten.

3.2.1 Massnahmen zur Strukturierung von Mathematikunterricht

Direkte und explizite Instruktion

Ansätze der direkten und der expliziten Instruktion gelten als äusserst wirksam für die Förderung bei Rechenschwäche (vgl. Kap. 3.1.1). In der Metastudie von Kroesbergen und van Luit (2003) wurden unter dem Begriff der direkten Instruktion mathematische Förderansätze subsumiert, die sich wie folgt beschreiben lassen:

> The direct instruction model has been used most frequently to teach task-specific strategies. Children are taught a specific action sequence and model that sequence in the given task context. Instruction is usually scripted and often is employed with groups of youngsters. This type of instruction is structured step by step to insure mastery before the student proceeds. Teacher instruction is gradually faded out, and practice and review are used to avoid precipitous loss once the instructional unit has been completed. (Goldman, 1989, S. 45)

Kern der direkten Instruktion sind also eine darbietende und direktive Stoff- bzw. Strategienvermittlung sowie angeleitetes Üben im Klassenverband, gefolgt von individuellem Üben und einem abschliessenden Rückblick auf das Gelernte (Hasselhorn & Gold, 2017). Dabei werden eine fortwährende Lernüberwachung und eine kontinuierliche Verstehensprüfung umgesetzt, indem die Lehrperson während der Stoffpräsentation bzw. während des Übens den Lernenden Verstehensfragen stellt und korrigierendes Feedback gibt (Hasselhorn & Gold, 2017). Ansätze der expliziten Instruktion sind ebenfalls stark strukturiert, direktiv und von der Lehrperson gelenkt und können daher nicht immer eindeutig von Ansätzen der direkten Instruktion unterschieden werden. Gemäss der Metastudie von Gersten et al. (2009, S. 1228) lassen sich die dort untersuchten Ansätze der expliziten Instruktion folgendermassen zusammenfassen:

> (a) The teacher demonstrated a step-by-step plan (strategy) for solving the problem; (b) the plan was problem specific and not a generic, heuristic guide for solving problems; and (c) students were actively encouraged to use the same procedure/step demonstrated by the teacher.

In Abgrenzung zur direkten Instruktion muss dazu angemerkt werden, dass bei der expliziten Instruktion vermehrt auch das Anregen der Schülerinnen und Schüler zum Verbalisieren eigener Denk- und Lösungswege und der gezielte Einsatz von Visualisierungen als wichtig erachtet wird (Gersten et al., 2009). Ersteres zeigt, dass die Rolle der Schülerinnen und Schüler bei der expliziten Instruktion eine aktivere ist als bei der direkten Instruktion. Montague (2011) weist zudem darauf hin, dass Ansätze der expliziten Instruktion auch in höherem Ausmass individualisierende Erklärungen und Unterstützungsmassnahmen durch die Lehrperson verlangen.

Dabei fokussieren insbesondere Ansätze der direkten Instruktion im Mathematikunterricht auf die Verbesserung des Faktenabrufs und die Vermittlung von Strategien, wobei das Vorgehen dabei eher als „drill and practice“ beschrieben wird (Montague, 2011, S. 298). So gibt die Lehrperson z.B. spezifische Lösungsstrategien Schritt für Schritt vor, und diese werden eins zu eins von den Schülerinnen und Schülern übernommen.

Kritisch anzumerken ist zudem, dass in den Metaanalysen von Kroesbergen und van Luit (2003) und Gersten et al. (2009) unklar bleibt, ob bei den dort untersuchten Förderansätzen auch längerfristige Effekte der Wirksamkeit zu finden sind. Dabei wäre es gerade bei Förderansätzen der direkten Instruktion, die bei Rechenfertigkeiten primär auf Auswendiglernen und intensives Einüben setzen (z.B. Fuchs et al., 2008; Powell et al., 2009b), wichtig zu untersuchen, ob diese nicht nur kurzfristig wirken. Fehlende Transfereffekte weisen zudem darauf hin, dass dabei nur sehr spezifisch die jeweils trainierten Fertigkeiten verbessert werden, die mathematische Entwicklung jedoch nicht umfassend gefördert wird (z.B. Fuchs et al., 2008). Daher wird unter Berücksichtigung von Schülerinnen und Schüler mit unterdurchschnittlichen mathematischen Leistungen die Gefahr betont, dass so gelernte Prozeduren und Lösungsstrategien zwar zwischenzeitlich gemäss Anweisungen angewendet werden können, aber kein tieferes Verständnis für diese aufgebaut wird, sodass sie z.B. Mühe haben, gelernte Strategien längerfristig anzuwenden und auf ähnliche, aber leicht abweichende Aufgaben zu übertragen (Hasselhorn & Gold, 2017; Montague, 2011). Es muss daher kritisch hinterfragt werden, ob sich solche stark strukturierten Förderansätze für den Aufbau konzeptuellen Verständnisses und Wissens im Fach Mathematik eignen (Pfister, Moser Opitz & Pauli, 2015). Da individuelle Strategien und Überlegungen zudem in Ansätzen der direkten Instruktion nicht oder kaum berücksichtigt werden, wird der Einsatz dieser Vorgehensweisen in heterogenen Lerngruppen erschwert und es muss angenommen werden, dass Leistungsunterschiede zwischen den Schülerinnen und Schülern eher verstärkt als gemindert werden (Hasselhorn & Gold, 2017; Montague, 2011). Dennoch bieten die direkte und die explizite Instruktion auch hilfreiche Hinweise für

Strukturierungsmassnahmen für stärker verstehensorientierte Förderansätze im Mathematikunterricht, wenn diese vermehrt individualisierend ausgestaltet werden. So sind gemäss Montague (2011) wichtige Gemeinsamkeiten stark strukturierter und effektiver Förderansätzen das Geben von gezielten Hinweisen, ein Lernen am Modell, eine verbale Begleitung von Lernprozessen, das Geben von Feedback sowie angeleitetes Üben.

Adaptive Lernbegleitung

Ansätze der adaptiven Lernbegleitung bieten ebenfalls wichtige Hinweise für Strukturierungsmassnahmen im Unterricht, die im Vergleich zu Massnahmen aus der direkten und expliziten Instruktion jedoch stärker die Chancen einer individualisierten Lernbegleitung nutzen und in grösserem Ausmass ein verstehensorientiertes Lernen ermöglichen, bei dem individuelle Denk- und Lösungsprozesse wichtiger Bestandteil des Unterrichts sind (Beck et al., 2008). Wie Parsons et al. (2018, 215f) in ihrer Übersichtsarbeit aufzeigen, wurden in Zusammenhang mit einer adaptiven Lernbegleitung in der Vergangenheit teilweise sehr unterschiedliche Begrifflichkeiten verwendet (u.a. „adaptive instruction", „adaptive teaching", „decision-making", „reflective teaching", „scaffolding", „teacher metacognition").

In diversen neueren Publikationen findet sich die Forderung, dass Lehrpersonen eine adaptive Lernbegleitung umsetzen sollen, um in leistungsheterogenen Klassen alle Schülerinnen und Schüler adäquat fördern zu können, wobei dies auch spezifisch für den Mathematikunterricht verlangt wird (z.B. Beck et al., 2008; C. Fischer, Kopmann, Rott, Veber & Zeinz, 2014; Hasselhorn & Gold, 2017; Leiss, 2010; Pfister et al., 2015). Dazu muss angemerkt werden, dass der Umgang mit Leistungsheterogenität bereits in den ersten Arbeiten, die sich mit dem Konzept einer adaptiven Lernbegleitung auseinandergesetzt haben, eine wichtige Rolle spielte (Parsons et al., 2018). Im Kern haben Ansätze der adaptiven Lernbegleitung als Gemeinsamkeit zum Ziel, dass alle Schülerinnen und Schüler trotz unterschiedlichem Vorwissen, Leistungsstand und unterschiedlichen motivationalen Bedingungen auf ihrem Niveau lernen können (Beck et al., 2008). Eine Lehrperson mit hoher adaptiver Lehrkompetenz zeichnet sich entsprechend dadurch aus, dass sie...

> […] es schafft, das Lehr-Lern-Geschehen unter bestmöglicher Berücksichtigung
> - der inhaltlichen Anforderungen des Unterrichtsinhaltes (Sachkompetenz),
> - der Vielfalt der Wissens- und Lernvoraussetzungen und der Lernverläufe der Schülerinnen und Schüler sowie der situativen Aspekte des Lernens (diagnostische Kompetenz),

> - der Möglichkeiten und Chancen der didaktischen Gestaltung der Lernsituation (didaktische Kompetenz),
> - der pädagogischen Maßnahmen zu Steuerung, Führung und Begleitung einer Schülergruppe oder Klasse (Klassenmanagement)
>
> erfolgreich zu orchestrieren […]. (Beck et al., 2008, S. 37)

Adaptive Kompetenz umfasst somit einerseits die Unterrichtsvorbereitung und -planung, andererseits ist damit aber auch ein Antizipieren des aktuellen Geschehens im Unterricht gemeint, da die Lehrperson ihr Handeln ständig den aktuellen Anforderungen der Lernenden anpassen muss (Beck et al., 2008; Parsons et al., 2018; Rogalla & Vogt, 2008). Für beide Aspekte benötigen Lehrpersonen Sachkompetenz, diagnostische Kompetenz, didaktische Kompetenz und Kompetenzen bezüglich der Klassenführung (Beck et al., 2008).

Häufig findet sich dabei ein Bezug zur soziokulturellen Lerntheorie von Vygotsky (Parsons et al., 2018). Gemäss Leiss (2010, S. 203) kann unter einer adaptiven Lernbegleitung im Sinne von Vygotsky (1978) „[…] die optimale Passung der Lehrerhandlungen in Bezug auf die individuellen, sozialen und kognitiven Voraussetzungen der Lernenden verstanden werden, um den Lernenden so das Erreichen der so genannten 'zone of proximal development' zu ermöglichen“. Die Zone der nächsten Entwicklung wird von Vygotsky folgendermassen definiert: „It is the distance between the actual developmental level as determined by individual problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, S. 86). Demnach gibt es beim Lernen eine Distanz zwischen dem, was ein Kind allein, und dem, was ein Kind mit Unterstützung durch eine erfahrenere Person lösen kann. Lernen ist gemäss Vygotsky (1978) daher am effektivsten, wenn es in der Zone der nächsten Entwicklung stattfindet. Vygotsky geht in seiner soziokulturellen Lerntheorie davon aus, dass Lernen immer in einem sozialen Kontext stattfindet und dass Lernen ein Prozess der Internalisation ist, für die eine aktive Auseinandersetzung in einem sozialen Kontext notwendig ist (Vygotsky, 1978). Im schulischen Kontext bedeutet dies, dass Kinder in der Schule auf Interaktionen untereinander sowie auf Interaktionen mit der Lehrperson angewiesen sind, was im nächsten Kapitel noch genauer ausgeführt wird.

Wie Pfister (2016) darlegt, besteht bei der adaptiven Lernbegleitung konzeptuell auch eine grosse Nähe zum Scaffolding-Ansatz, wobei es in beiden Ansätzen das Ziel ist, Lernprozesse optimal zu unterstützen. In der ursprünglichen Konzeptualisierung von Wood, Bruner und Ross (1976) wurde der Begriff Scaffolding für Eins-zu-eins-Unterstützungssituationen verwendet.

Beim Scaffolding geht es darum, dass eine erfahrenere Person eine weniger erfahrene Person dabei unterstützt, ein Problem oder eine Aufgabe zu bearbeiten, die ohne Unterstützung noch nicht allein gelöst werden kann (Wood et al., 1976). Die erfahrenere Person strukturiert und modelliert dabei die Lernsituation so, dass die weniger erfahrene Person die Aufgabe lösen kann (Wood et al., 1976). Scaffolding kann daher als höchst adaptive Form der Lernbegleitung bezeichnet werden, bei der ein hohes Ausmass an Strukturierung vorliegt. Als gemeinsamer Kern verschiedener Definitionen von Scaffolding werden drei Schlüsselmerkmale genannt: „contingency", „fading" und „transfer of responsibility" (van de Pol, Volman & Beishuizen, 2010). Es kann demnach von Scaffolding gesprochen werden, wenn die Unterstützung an das aktuelle Fähigkeits- und Kompetenzniveau angepasst ist („contingency"), wenn die Unterstützung im Verlaufe der Zeit unter Berücksichtigung der aktuellen Fähigkeiten und Kompetenzen graduell abnimmt („fading") und wenn der oder die Lernende im Verlaufe des Unterstützungsprozesses immer stärker die Verantwortung für den Problemlöseprozess übernimmt („transfer of responsibility"; van de Pol et al., 2010). Weiter werden sechs Scaffolding-Strategien genannt, mit denen Lernprozesse adaptiv unterstützt werden können (vgl. van de Pol et al., 2010, S. 277):

- *Feeding back*: Lernenden Feedback geben zum Lernprozess,
- *Hints*: Gezielte Hinweise geben, die für das Voranschreiten des Lösungsprozesses notwendig sind,
- *Instructing*: Anweisungen an Lernende, was, wie und warum etwas gemacht werden muss,
- *Explaining*: Das Geben von zusätzlichen Informationen und Klärungen,
- *Modeling*: Vorzeigen von Handlungen und Problemlöseschritten, damit diese imitiert werden können,
- *Questioning*: Stellen von Fragen, die zum Versprachlichen von und Denken über Sachverhalte anregen.

Des Weiteren lassen sich nach Pfister et al. (2015) für den Mathematikunterricht fünf Facetten von Scaffolding herausarbeiten, die sich für eine adaptive Lernbegleitung eignen. Zentrale Facetten von Scaffolding im Mathematikunterricht sind demnach:

- *Kognitive Aktivierung* durch das Stellen von herausfordernden Aufgaben und Fragen,
- *Anregen von Diskursen,* um einen gemeinsamen Austausch und eine Reflexion über das Gelernte zu ermöglichen,

- *Produktiver Umgang mit Fehlern,* um individuelle Lernprozesse geeignet zu unterstützen, indem an den Schwierigkeiten der Schülerinnen und Schüler angesetzt wird,
- *Zielorientierung,* indem Hinweise gegeben bzw. zentrale Erkenntnisse rekapituliert werden, um die Schülerinnen und Schüler zum Kern der Sache zu führen,
- *Produktiver Einsatz von Arbeitsmitteln,* um Einsicht in mathematische Strukturen zu ermöglichen und individuelle Lernprozesse zu unterstützen (vgl. Pfister et al., 2015, S. 1084).

Vergleicht man die Strategien für eine individuelle Lernunterstützung von van de Pol et al. (2010) mit den bereits erwähnten Merkmalen wirksamer, stark strukturierter Instruktionsansätze von Montague (2011), können einige Parallelen gefunden werden (z.B. Feedback, gezielte Hinweise und Modellieren). Im Vergleich zur direkten Instruktion geht es bei der adaptiven Lernbegleitung jedoch darum, dass Strukturierungsmassnahmen stärker an individuellen Lern- und Verstehensprozessen ansetzen, sodass diese Herangehensweise für heterogene Lerngruppen als besser geeignet angesehen wird (vgl. Leiss, 2010). Für die Gestaltung der unterrichtsintegrierten Mathematikförderung, die in dieser Arbeit untersucht werden soll, werden daher die in diesem Kapitel diskutierten Massnahmen zur individuellen Lernbegleitung und zur Strukturierung von Lernprozessen aus dem Fachdiskurs zur Adaptivität und zu Scaffolding berücksichtigt und durch Erkenntnisse zu Strukturierungsmassnahmen aus Ansätzen der direkten und expliziten Instruktion in geeigneter Weise ergänzt.

3.2.2 Bedeutung gemeinsamer Lernsituationen

Zusätzlich zu einer adaptiven Lernbegleitung werden innere Differenzierungs- und Individualisierungsmassnahmen als weitere zentrale Instrumente angesehen, um in heterogenen Schulklassen alle Schülerinnen und Schüler adäquat auf ihrem Leistungs- und Lernniveau zu fördern (z.B. Hugener, Krammer & Pauli, 2008; Leuders & Prediger, 2016; Scherer & Moser Opitz, 2010). Dabei besteht jedoch die Gefahr, dass Differenzierungs- und Individualisierungsmassnahmen zu einer sozialen Vereinzelung führen können, wenn sie, wie dies häufig der Fall ist, primär in Form von einem Abarbeiten individueller Arbeitspläne umgesetzt werden (Moser Opitz, 2009; Vorstand des Arbeitskreises Schweiz-Liechtenstein der GDM, 2013). Die gleiche Problematik stellt sich auch bei Fördersettings, in denen auf eine Einzelförderung gesetzt wird. Diese Vereinzelung wird aus inklusionspädagogischer und integrationspädagogischer, aber auch aus (fach)didaktischer und empirischer Sichtweise als problematisch angesehen (Stöckli, Moser Opitz, Pfister & Reusser, 2014). Um dem entgegenzuwirken, wird als zentral erachtet, dass in heterogenen Lerngruppen gemeinsame Lernsituationen im Unterricht ermöglicht werden. Diese werden von verschiedenen Autorinnen und Autoren unterschiedlich definiert und begründet.

Bei Feuser (1989, 2013) ist das Prinzip der Kooperation am gemeinsamen Gegenstand zentraler Eckpunkt der entwicklungslogischen Didaktik. Diese versteht er als allgemeine Pädagogik für alle Schülerinnen und Schüler, unabhängig vom jeweiligen Leistungsniveau, verbunden mit der Forderung nach einer allgemeinen Schule, die ohne Selektions- und Segregationsmechanismen auskommt (Feuser, 1989). Ein hohes Mass an innerer Differenzierung und Individualisierung ist für eine solche Schule unerlässlich, um allen Kindern optimale Lern- und Entwicklungsmöglichkeiten zu gewährleisten. Eine rein räumliche Zusammenführung bei einer Beibehaltung individueller Curricula für verschiedene Schülerinnen- und Schülergruppen wird ausdrücklich abgelehnt (Feuser, 1989). Vielmehr wird ein soziales Miteinander in Form einer Kooperation am gemeinsamen Gegenstand als notwendige Voraussetzung für Lern- und Entwicklungsprozesse angesehen: „[…] Erkenntnisbildung [setzt] in gleicher Weise handelndes Wirken in der Welt wie wirkende zwischenmenschliche Beziehungen voraus. *Der Mensch erschliesst sich die Dinge durch den Menschen und sich den Menschen über die Dinge – in der gemeinsamen Kooperation*“ (Feuser, 2013, S. 28).

Als zur Umsetzung geeignet werden offene Unterrichtsgestaltungen z.B. in Form eines fächerübergreifenden Projektunterrichts vorgeschlagen, der idealerweise in jahrgangsübergreifenden Klassen stattfinden sollte (Feuser, 1989). Dabei muss Unterricht didaktisch von unten nach oben geplant werden. D.h., es muss herausgearbeitet werden, welche elementaren und fundamentalen Ideen ausgehend vom Entwicklungsniveau und von den Handlungs- und Denkkompetenzen der leistungsschwächsten Schülerinnen und Schülern an einem gemeinsamen Gegenstand erarbeitet werden können (Feuser, 1989). Hier ist wichtig, dass der gemeinsame Gegenstand „[…] *nicht das materiell Fassbare* [*ist*], das letztlich in der Hand des Schülers zum Lerngegenstand wird, *sondern der zentrale Prozess*, der hinter den Dingen und beobachtbaren Erscheinungen steht und sie hervorbringt“ (Feuser, 1989, S. 32). Demnach ist der gemeinsame Gegenstand nicht eine bestimmte zu bearbeitende Aufgabe oder ein bestimmtes Material, das von allen zur Lösung der Aufgabe herangezogen wird, sondern das Fundamentale und Elementare einer Idee oder eines Inhaltes.

Wocken (1998) anerkennt die Wichtigkeit einer Kooperation am gemeinsamen Gegenstand, kritisiert jedoch, dass es viele weitere Formen von gemeinsamen Lernsituationen gebe, die ebenfalls im integrativen Sinne eine gemeinsamkeitsstiftende Funktion hätten (Wocken, 1998). Von grosser Bedeutung seien diesbezüglich kommunikative Lernsituationen, bei denen es sich um reine Interaktionsprozesse handle, ohne dass eine der beiden beteiligten Personen bestimmte Lernziele verfolge. So argumentiert der Autor, dass ein Gespräch auf dem Schulweg oder in der Pause viel stärker zu einem echten Miteinander verschiedener Kindern führen könne als eine didaktisch speziell dafür arrangierte Situation während des Unterrichts. Gemäss Wocken (1998) ermöglichen aber auch koexistente Lernsituationen, wie sie von (Feuser, 1989, 2013) ausdrücklich abgelehnt werden, gemeinsamkeitsstiftende Momente im Unterricht. Hier handelt es sich um Lernsituationen, bei denen ein rein räumliches Beieinandersein im Vordergrund steht, ohne dass es zu nennenswerter sozialer Interaktion kommt. Weitere Formen von gemeinsamen Lernsituationen nach Wocken (1998) sind subsidiäre Lernsituationen, bei denen für die Zielerreichung die Unterstützung durch eine andere Person notwendig ist, und kooperative Lernsituationen, bei denen beide für die eigene Zielerreichung auf die Handlungen der anderen Person angewiesen sind.

Freudenthal (1974) wiederum plädiert aus mathematikdidaktischer Perspektive für gemeinsame Lernsituationen für Schülerinnen und Schüler mit unterschiedlichen Leistungsniveaus. Ein zentrales Argument hierfür ist, dass die natürliche Struktur mathematischer Lernprozesse eine Logik aufweist, die sich per se dafür eignet, dass verschiedene Kinder am gleichen Gegenstand auf verschiedenen Stufen des Verständnisses arbeiten können (Freudenthal, 1974). So schreibt er:

> In einer Gruppe sollen die Schüler zusammen, aber jeder auf der ihm gemäßen Stufe, am gleichen Gegenstande arbeiten, und diese Zusammenarbeit soll es sowohl denen auf niedriger Stufe wie denen auf höherer Stufe ermöglichen, ihre Stufe zu erhöhen, denen auf niedrigerer Stufe, weil sie sich auf die höhere Stufe orientieren können, denen auf höherer Stufe, weil die Sicht auf niedrigere Stufen ihnen neue Einsichten verschafft. (Freudenthal, 1974, S. 167)

Einen wichtigen Stellenwert in gemeinsamen Lernsituationen im Mathematikunterricht hat gemäss Freudenthal (1974) die Versprachlichung mathematischer Sachverhalte, da so Erkenntnisse verschiedener Schülerinnen und Schüler auf unterschiedlichen Stufen des Verständnisses miteinander in Bezug gesetzt werden können.

Für die Bedeutsamkeit gemeinsamer Lernsituationen sprechen auch einige empirische Erkenntnisse. In der Untersuchung von Wiener und Tardif (2004) zeigte sich, dass stärker integrative Settings sich u.a. positiv auf Aspekte der sozialen Partizipation von Schülerinnen und Schülern mit leichten bis gravierenden Formen von Lernschwierigkeiten ausgewirkt haben. Wurde eine Förderung innerhalb des regulären Klassenraumes („in-class support“) im Gegensatz zu einer

Förderung in separaten Räumen („ressource room support") umgesetzt, ergaben sich für Schülerinnen und Schüler mit leichten bis mittleren Lernschwierigkeiten höhere Werte bezüglich der sozialen Akzeptanz, ein positiveres Selbstkonzept der eigenen mathematischen Leistungsfähigkeit und seltener auffällige Verhaltensweisen. Auch für Inklusionsklassen zeigten sich positive Effekte für Schülerinnen und Schüler mit gravierenden Formen von Lernschwierigkeiten im Hinblick auf die selbst berichtete Qualität von Freundschaften, die eigene Einschätzung bezüglich Einsamkeit und die Häufigkeit von auffälligem Verhalten (Wiener & Tardif, 2004). Eine Zusammenstellung relevanter Studien, die Interventionsmassnahmen im Bereich der sozialen Partizipation umgesetzt haben, konnte zudem die Wichtigkeit akademischer Gruppenaktivitäten u.a. in Form kooperativer Lernsettings bestätigen, wobei sich positive Effekte für schulische Leistungen, soziale Interaktionen und die soziale Akzeptanz von Kindern mit Lern- und Verhaltensschwierigkeiten ergaben (Garrote, Sermier Dessemontet & Moser Opitz, 2017).

Zusammenfassend lässt sich festhalten, dass gemeinsame Lernsituationen insbesondere auch unter Berücksichtigung heterogener Lerngruppen aus verschiedenen Gründen als wichtig erachtet und deren Bedeutsamkeit auch empirisch bestätigt werden konnte. So wird einerseits angenommen, dass gemeinsame Lernsituationen einen gemeinsamkeitsstiftenden Effekt haben und somit für eine inklusive bzw. integrative Schule wichtig sind. Andererseits sind gemeinsame Lernsituationen auch für mathematische Lernprozesse und die mathematische Entwicklung an sich bedeutsam. Um einen produktiven, fachbezogenen Austausch im Mathematikunterricht zu ermöglichen, müssen Lehrpersonen jedoch mithilfe geeigneter Strategien die Schülerinnen und Schüler zu einem solchen Austausch anregen (vgl. Kap. 3.3). Angesichts von Klassen mit einem heterogenen Lern- und Leistungsniveau der Schülerinnen und Schüler muss dabei beachtet werden, dass dafür ein gemeinsamer Gegenstand, auf den sich der Austausch beziehen kann, als unabdingbar erachtet wird (Feuser, 2013). In der in dieser Arbeit zu untersuchenden Mathematikförderung werden daher gemeinsame Lernsituationen geschaffen, in denen im Klassenverband ein kommunikativer Austausch stattfinden kann.

3.3 Mathematikdidaktische Überlegungen zur Förderung rechenschwacher Schülerinnen und Schüler

Zusätzlich zu den bisher genannten empirischen und didaktischen Hinweisen bezüglich einer effektiven Förderung von Schülerinnen und Schülern mit einer Rechenschwäche liefert auch die Mathematikdidaktik wertvolle Anhaltspunkte zur Gestaltung von förderlichen Lernangeboten im Mathematikunterricht unter spezieller Berücksichtigung von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen. Dabei lassen sich inhaltsübergreifende

Überlegungen (Fokussierung auf zentrale mathematische Inhalte, produktive Übungsformate, Einsatz von Arbeitsmitteln und Veranschaulichungen) von inhaltsspezifischen Aspekten (zählendes Rechnen, dezimales Stellenwertsystem, ordinaler Zahlaspekt, Mathematisieren und Operationsverständnis, halbschriftliche Rechenstrategien) unterscheiden, so dass diese im Folgenden separat thematisiert werden.

3.3.1 Inhaltsübergreifende Überlegungen

Fokussierung auf zentrale mathematische Inhalte

Als unbestritten gilt, dass eine mathematische Förderung auf zentrale mathematische Inhalte fokussieren und auf einen systematischen Kompetenzaufbau unter Berücksichtigung mathematischer Entwicklungsmodelle abzielen muss (Koch, 2008; Krajewski, 2008b; W. Schneider et al., 2013; von Aster et al., 2005). Eine Mathematikförderung muss demnach direkt an mathematischen Inhalten Verständnis und Einsicht aufbauen, anstatt bei basalen Defiziten z.B. in der Wahrnehmung und der Motorik anzusetzen (Krajewski, 2008b; von Aster et al., 2005). Diese Aussage mag banal klingen, ist jedoch eine Reaktion auf die lange Zeit verbreitete Praxis in der Sonderpädagogik, mithilfe von Motorik- und Wahrnehmungstrainings eine Verbesserung von Mathematikleistungen zu erreichen (vgl. z.B. Milz, 1993), was sich jedoch nicht als zielführend erwiesen hat (Grünke, 2006; vgl. auch Gold, 2016). Bei der Frage, welche Inhalte gefördert werden sollen, können empirisch fundierte mathematische Entwicklungsmodelle (vgl. Kap. 2.2) hilfreiche Hinweise dazu bieten, welche Kompetenzen für die weitere mathematische Entwicklung von besonderer Bedeutung sind (Krajewski, 2008b). Da diese Entwicklungsmodelle jedoch primär die frühe mathematische Entwicklung abdecken, sollten zusätzlich auch empirische Erkenntnisse zu inhaltlichen Schwierigkeiten älterer rechenschwacher Schülerinnen und Schüler berücksichtigt werden (vgl. Kap. 2.4). Wie Moser Opitz und Freesemann (2012) schreiben, muss mit rechenschwachen Schülerinnen und Schülern jedoch nicht der gesamte Lernstoff behandelt bzw. aufgearbeitet werden, sondern es muss eine Fokussierung auf zentrale Inhalte, die für die weitere mathematische Entwicklung bedeutsam sind, erfolgen. Moser Opitz (2013) kommt auf Basis empirischer Erkenntnisse zu inhaltlichen Schwierigkeiten rechenschwacher Kinder sowie auf Grund von fachlichen und mathematikdidaktischen Überlegungen zum Schluss, dass folgende Inhalte für die ersten vier Klassen der Primarschule als zentral zu erachten sind:

- Erwerb der Zählkompetenz
- Verständnis Dezimalsystem
- Teil-Ganze-Beziehungen (z. B. Zahlzerlegungen, Ergänzen, Verdoppeln)

- Inverses Prinzip der Addition und Subtraktion
- Konzeptuelles Verständnis der Multiplikation
- Einsicht in Rechengesetze
- Multiplikationsaufgaben mit den Faktoren 2 und 5 (→ verdoppeln, halbieren)
- Halbschriftliches Rechnen
- Schriftliches Rechnen: Einsicht erwerben
- Problemlösen: Situations- und Problemlösemodell aufbauen. (Moser Opitz, 2013, S. 138)

Diese Inhalte werden als mathematischer Basisstoff bezeichnet, und es wird empfohlen bei einer Förderung für rechenschwache Schülerinnen und Schüler inhaltlich an diesem Basisstoff anzusetzen (Moser Opitz, 2013), wobei in tieferen Klassenstufen gewisse Inhalte noch weggelassen werden können (z.B. schriftliche Rechenverfahren).

Produktive Übungsformate

Üben ist ein wesentlicher Bestandteil des mathematischen Kompetenzaufbaus und nimmt im Mathematikunterricht einen grossen Stellenwert ein. Als problematisch erachtet werden jedoch Formen des Übens, die primär auf eine Automatisierung im Sinne von Auswendiglernen abzielen oder eine mechanische Bearbeitung beliebig aneinandergereihter Aufgaben beinhalten (Krauthausen & Scherer, 2007; Scherer & Moser Opitz, 2010). Insbesondere für rechenschwache Schülerinnen und Schüler wird für eine gewinnbringende Automatisierung von Faktenwissen gefordert, dass zunächst Einsicht und Verständnis für den Inhalt aufgebaut werden, gefolgt von einschleifenden Übungsformaten, in denen das Gelernte in strukturierter Form automatisiert wird (Krauthausen & Scherer, 2007).

Wittmann (1992, S. 177) unterscheidet diesbezüglich zwischen einführenden Übungen, bei denen „[…] typische Aufgaben betrachtet [werden], ggf. nur wenige, und die zu lernenden Wissenselemente bzw. Fertigkeiten werden an ihnen erarbeitet“ und Übungsaufgaben im engeren Sinne, bei denen zu erwerbendes Wissen bzw. Kompetenzen an einer grösseren Anzahl vergleichbarer Aufgaben eingeübt werden. Ein Übungsformat, das sich für Letzteres besonders gut eignet, sind operativ strukturierte Übungsformate im Sinne von Wittmann (1992) (Scherer & Moser Opitz, 2010). Dabei soll eine systematische Variation in den Aufgaben dazu führen, dass Gesetzmässigkeiten erkannt werden können. Damit sind also Aufgabenserien gemeint, in denen die einzelnen Aufgaben, die Lösungswege oder Ergebnisse untereinander in einem systematischen Bezug zueinander stehen, sodass bei der Bearbeitung der Aufgaben Strukturzusammenhänge ersichtlich werden (Wittmann, 1992; vgl. Abbildung 4). Wie Moser Opitz

(2008b) mit Rückgriff auf Scherer (1995) darlegt, können solche Aufgaben für schwache Rechnerinnen und Rechner zusätzlich zum Anregen von Verständnis für und von Einsicht in Strukturzusammenhänge auch eine Entlastung des Gedächtnisses bieten. „Erfahrungen zeigen zudem, dass lernschwache Kinder häufig Spass an diesen operativen Beziehungen, den „Mustern“ haben, sodass auch die motivationale Komponente berücksichtigt werden kann […]“ (Scherer, 2009, S. 443). Für einen nachhaltigen Verständnisaufbau ist wichtig, dass solche Aufgabenformate insbesondere mit Blick auf rechenschwache Schülerinnen und Schüler immer wieder angeboten werden (Scherer & Moser Opitz, 2010).

7 + 31 =	17 + 31 =
17 + 31 =	45 + 12 =
27 + 31 =	52 + 36 =
37 + 31 =	23 + 25 =

Abbildung 4: Beispiel für eine operativ strukturierte Aufgabenserie (links) im Vergleich zu einer unstrukturierten Aufgabenserie (rechts) nach Scherer und Moser Opitz (2010, S. 64)

Ein weiteres sinnvolles Übungsformat können offene Aufgaben darstellen (Scherer, 2009; Scherer & Moser Opitz, 2010). Offene Aufgaben sind Aufgaben, in denen die Schülerinnen und Schüler selber den Schwierigkeitsgrad, auf dem sie eine Aufgabe lösen, bestimmen können (z.B. „Finde so viele Rechnungen wie möglich, bei welchen das Ergebnis 100 lautet“; Scherer & Moser Opitz, 2010). Diese können im Sinne einer inneren Differenzierung sinnvoll für den Unterricht in leistungsheterogenen Schulklassen sein, da sie allen Kindern eine Bearbeitung der Aufgaben auf ihrem aktuellen Lernniveau ermöglichen (Hußmann & Prediger, 2007; Scherer & Moser Opitz, 2010).

Einsatz von geeigneten Veranschaulichungen und Arbeitsmitteln

Der Einsatz von geeigneten Veranschaulichungen und Arbeitsmitteln wird ebenfalls als zentraler Bestanteil einer mathematischen Förderung für rechenschwache Schülerinnen und Schüler angesehen (Koch, 2008; Krajewski, 2008b; W. Schneider et al., 2013; Van de Walle, 2007; von Aster et al., 2005). Damit soll der Aufbau „[…] klarer, tragfähiger mentaler Vorstellungsbilder […]“ ermöglicht werden (Krauthausen & Scherer, 2007, S. 246). Zudem kann die Verwendung von Veranschaulichungen und Arbeitsmitteln helfen, das Arbeitsgedächtnis rechenschwacher Schülerinnen und Schüler zu entlasten (Scherer & Moser Opitz, 2010; W. Schneider et al., 2013; vgl. auch Kap. 2.3.1).

Veranschaulichungen dienen in erster Linie dazu, bestimmte mathematische Sachverhalte und Ideen zu erläutern und darzustellen (Krauthausen & Scherer, 2007). Arbeitsmittel können im Gegensatz dazu auch als Hilfsmittel für das Lösen von mathematischen Operationen verwendet werden, indem an diesen direkt mathematische Handlungen vollzogen werden (Krauthausen & Scherer, 2007).[10] Je nach verwendetem Kontext können sich aber auch Überschneidungen ergeben, wenn beispielsweise ein Arbeitsmittel zur Veranschaulichung genutzt wird und umgekehrt. Beispiele für Materialien, die in erster Linie als Veranschaulichungen eingesetzt werden, sind der Zahlenstrahl und die Hunderterkette, die den ordinalen Zahlaspekt betonen (Scherer & Moser Opitz, 2010). Arbeitsmittel können zudem nach ihrem Strukturierungsgrad unterschieden werden (Radatz, Schipper, Dröge & Ebeling, 1996). Unstrukturierte Materialien wie lose Plättchen haben den Vorteil, dass eine flexible Darstellung kleiner Anzahlen möglich ist und die Plättchen unterschiedlich gebündelt werden können (Scherer & Moser Opitz, 2010). Ein Nachteil ist jedoch, dass ab einer Anzahl von etwa vier bis fünf keine simultane Erfassung mehr möglich ist, die Schülerinnen und Schüler dazu verleitet werden können, zählend zu rechnen, und es bei grossen Anzahlen schwierig ist, die Übersicht über das Material zu behalten (Scherer & Moser Opitz, 2010). Strukturierte Materialien mit festen Einheiten wie z.B. das Dienes-Material weisen hingegen eine Zehnerstruktur (oder Fünferstruktur) auf, ermöglichen eine strukturierte (quasi-simultane) Anzahlerfassung und eignen sich für die Darstellung grösserer Anzahlen (Radatz et al., 1996; Scherer & Moser Opitz, 2010). Das Dienes-Material, das den kardinalen Zahlaspekt betont und somit helfen kann, eine Vorstellung der Mächtigkeit von Zahlen aufzubauen, hat zudem den Vorteil, dass damit auch das Bündelungs- und Entbündelungsprinzip in Zusammenhang mit der dezimalen Stellenwertschreibweise gut veranschaulicht wird (Scherer & Moser Opitz, 2010; Schmassmann & Moser Opitz, 2008a). Strukturierte Materialien mit flexiblen Einheiten (z.B. Plättchen in Kombination mit Hunderterfeld, Rechenrahmen) ermöglichen wiederum sowohl eine strukturierte (quasi-simultane) Anzahlerfassung als auch das Operieren mit flexiblen Einheiten (Scherer & Moser Opitz, 2010).

Da sich die verschiedenen Veranschaulichungen und Arbeitsmittel für unterschiedliche Zwecke eignen, müssen diese auf Basis mathematikdidaktischer Überlegungen sinnvoll ausgewählt werden (Krauthausen & Scherer, 2007; Scherer & Moser Opitz, 2010; Van de Walle, 2007). Zudem müssen Veranschaulichungen und Arbeitsmittel insbesondere unter Berücksichtigung rechenschwacher Schülerinnen und Schüler explizit eingeführt und deren Struktur und Ver-

[10] Krauthausen und Scherer (2007, S. 242) benutzen in ihrem Buch die Begriffe „Veranschaulichungsmittel“ und „Anschauungsmittel“, verwenden die Begriffe „Veranschaulichungen“ und „Arbeitsmittel“ jedoch in der Kapitelüberschrift, sodass hier unter Berücksichtigung der Begriffsverwendung bei Moser Opitz (2013, S. 75) letztere Begriffe synonym für die Begriffe „Veranschaulichungsmittel“ und „Anschauungsmittel“ verwendet werden.

wendungszweck mit den Schülerinnen und Schülern erkundet und erarbeitet werden (Krauthausen & Scherer, 2007; Scherer & Moser Opitz, 2010). Hinweise, wie dies konkret gemacht werden kann, werden bei den inhaltsspezifischen Überlegungen am Beispiel des Dienes-Materials und der Hunderterkette noch ausgeführt. Weiter müssen „[…] die Arbeitsmittel und Veranschaulichungen so eingesetzt werden, dass Vorstellung aufgebaut werden kann. Das kann durch geeignete Aufgabenstellungen erreicht werden, indem die Sprache als handlungsbegleitendes Mittel eingesetzt wird“ (Scherer & Moser Opitz, 2010, S. 199). Es ist also von grosser Bedeutung, das Material nicht nur handelnd einzusetzen, sondern die mathematischen Handlungen auch zu versprachlichen und gemeinsam zu reflektieren. „Insbesondere für lernschwache Schülerinnen und Schüler ist nicht immer klar, was eine soeben vorgenommene Handlung […] mit der dazugehörigen symbolischen Notationsform zu tun hat. Durch die Sprache kann die Verbindung zwischen der Handlung und den Symbolen explizit hergestellt und damit sichtbar gemacht werden […]“ (Scherer & Moser Opitz, 2010, S. 86). Die Schülerinnen und Schüler sollen daher ermuntert werden, mathematische Handlungen, die sie an Materialien durchführen, zu protokollieren, zeichnerisch darzustellen und zu versprachlichen (Schmassmann & Moser Opitz, 2008b).

Es wird weiter empfohlen, besonders bei rechenschwachen Kindern nicht zu viele verschiedene Materialien auf einmal einzuführen, sondern sich auf einige wesentliche Materialien zu konzentrieren (Krauthausen & Scherer, 2007; Landerl & Kaufmann, 2008; Scherer & Moser Opitz, 2010). Eine Ablösung von der Verwendung von Veranschaulichungen und Hilfsmitteln soll gemäss Krauthausen und Scherer (2007, S. 256) nicht erzwungen werden, sondern mithilfe von „[…] Übungen, die zum Aufbau tragfähiger Bilder und des mentalen Operierens beitragen […]“ gezielt ermöglicht werden. Allerdings wird gefordert, dass rechenschwache Kinder Arbeitsmittel und Veranschaulichungen nicht als permanente „Krücke“ verwenden, wie dies teilweise im sonderpädagogischen Unterricht der Fall ist (Schmassmann & Moser Opitz, 2008b). Arbeitsmittel und Veranschaulichungen sollen jedoch „[…] immer wieder zur Auffrischung, als Erinnerungsstütze oder zum Erkunden und Erforschen neuer Zusammenhänge ganz selbstverständlich zur Verfügung stehen“ (Schmassmann & Moser Opitz, 2008b, S. 41). Hilfreiche Hinweise, wie das Übersetzen von äusseren Handlungen in innere Handlungen gefördert und so eine Ablösung von Arbeitsmitteln und Veranschaulichungen angeregt werden kann, finden sich im Vierphasenmodell zum Aufbau von Grundvorstellungen von Wartha und Schulz (2011), das in dieser Arbeit in Zusammenhang mit der Förderung von Mathematisierungskompetenzen bei den inhaltsspezifischen Überlegungen noch genauer erläutert wird.

3.3.2 Inhaltsspezifische Überlegungen

Zählendes Rechnen

Eine Ablösung vom zählenden Rechnen wird als wichtiger Schritt in der mathematischen Entwicklung angesehen (Häsel-Weide, 2016), da zählende Rechenstrategien zeitaufwendig, fehleranfällig und fürs Rechnen mit grossen Zahlen ungeeignet sind (Gaidoschik, 2012; Moser Opitz, 2013). Für eine Ablösung vom zählenden Rechnen werden verschiedene Massnahmen empfohlen (Häsel-Weide, 2016; Scherer & Moser Opitz, 2010):

- Förderung flexibler Zählkompetenzen
- Förderung von Einsicht in Teile-Ganzes-Beziehungen
- Förderung einer strukturierten Anzahlerfassung
- Förderung von Vorstellungen für Operationen

Flexible Zählkompetenzen können dadurch gefördert werden, dass das Zählen in Zweier-, Fünfer- und Zehnerschritten, das Zählen von verschiedenen Startzahlen und das Rückwärtszählen geübt werden (Häsel-Weide, 2016; Moser Opitz, 2013; Scherer & Moser Opitz, 2010; Schmassmann & Moser Opitz, 2008a). Dabei empfiehlt es sich, bei rechenschwachen Kindern Arbeitsmittel und Veranschaulichungen einzusetzen, um das Arbeitsgedächtnis zu entlasten (Scherer & Moser Opitz, 2010). Für den Erwerb einer flexiblen Zählkompetenz ist wichtig, dass diese eine Fünfer- und/oder eine Zehnerstrukturierung aufweisen, die dabei hilft, vom Zählen in Einerschritten hin zum Zählen in grösseren Schritten zu gelangen (Häsel-Weide, 2016; Scherer & Moser Opitz, 2010). Dazu eignet sich z.B. eine bildliche Darstellung der Hunderterkette, bei der nur die Fünfer- und Zehnerzahlen beschriftet sind.

Als weiterer wichtiger Aspekt für die Ablösung vom zählenden Rechnen wird die Einsicht in Teile-Ganzes-Beziehungen angesehen, die durch Aufgaben zu Zahlzerlegungen sowie durch Verdoppelungs- und Halbierungsaufgaben geübt und längerfristig automatisiert werden sollen (Häsel-Weide, 2016; Scherer & Moser Opitz, 2010). Hierfür können operativ strukturierte Übungsformate im Sinne von Wittmann (1992), bei denen Zahlen systematisch variiert und von einfachen Kernaufgaben schwierigere Aufgaben abgeleitet werden können, hilfreich sein (Häsel-Weide, 2016; Scherer & Moser Opitz, 2010). Zudem wird empfohlen, die strukturierte Anzahlerfassung zu üben (insbesondere die Fünferstruktur), sodass kleine Mengen nicht mehr zählend erfasst werden müssen (Gerster, 1996; Krauthausen, 1995; Scherer & Moser Opitz, 2010). Dabei sollen Kinder selber verschiedene Strukturierungen von kleinen Anzahlen erproben, damit sie Strukturierungen finden, die sie auf einen Blick erfassen können (Scherer & Moser Opitz, 2010). Wichtig ist aber auch, dass die Strukturierungen reflektiert werden, indem diese versprachlicht oder durch Handlungen bzw. Zeichnungen veranschaulicht werden. Geeignete Möglichkeiten für die Förderung von Vorstellungen der vier Grundoperationen werden in einem separaten Abschnitt (*Mathematisieren und Operationsverständnis*, S. 72) erläutert.

Dezimales Stellenwertsystem

Eines der zentralsten Konzepte des Arithmetikunterrichts der Primarschule ist das dezimale Stellenwertsystem, wobei die Schülerinnen und Schüler nebst der Stellenwertschreibweise insbesondere auch die Prinzipien des Bündelns und Entbündelns verstehen müssen (Krauthausen & Scherer, 2007; Müller & Wittmann, 1984; Scherer & Moser Opitz, 2010; Van de Walle, 2007).

> Wenn Schülerinnen und Schüler nicht bzw. nur teilweise über Einsicht ins dezimale Stellenwertsystem verfügen, fehlen wichtige Voraussetzungen für erfolgreiche arithmetische Lernprozesse: das Verständnis von Zahlen und damit verbunden die Basis für den Erwerb der Grundoperationen […]; das Verständnis von grossen Zahlen, Zahlvorstellungen, Dezimalzahlen und Größen; die Basis für das Schätzen, Überschlagen und das Runden von Zahlen […]. (Scherer & Moser Opitz, 2010, S. 130)

Das Prinzip des Entbündelns ist zudem eine zentrale Voraussetzung für das Verständnis der Subtraktionen mit Übergängen und für das Rückwärtszählen (Scherer & Moser Opitz, 2010). Krauthausen und Scherer (2007, S. 16) weisen darauf hin, dass das dezimale Zahlensystem ein so selbstverständlicher Teil der Mathematik ist, sodass es notwendig ist, „die dahinterstehenden grundlegenden Prinzipien wieder in den Bewusstseinshorizont zu rufen und mit wirklichem Verstehen zu füllen“. Mit der Erweiterung des Zahlenraums auf tausend in der dritten Klasse eignet sich diese Klassenstufe daher besonders gut, um zentrale Aspekte und bereits erworbene Kenntnisse des dezimalen Stellenwertsystems nochmals zu erarbeiten und zu vertiefen, da z.B. das Prinzip der fortgesetzten Bündelung bei grossen Zahlen seine Bedeutsamkeit erst richtig entfalten kann (Scherer & Moser Opitz, 2010).

Für die Erarbeitung eines tiefgehenden Verständnisses des dezimalen Stellenwertsystems wird empfohlen, zunächst das Bündelungsprinzip zu erarbeiten (Freesemann, 2014; Scherer & Moser Opitz, 2010). Van de Walle (2007) schlägt hierfür vor, mit dem Dienes-Material zu arbeiten. Dabei müssen die Schülerinnen und Schüler die Gelegenheit haben, das Material zunächst zu erkunden. Es muss mit den Schülerinnen und Schülern explizit erarbeitet werden, dass ein Zehnerstab aus zehn Einerwürfeln, dass eine Hunderterplatte aus zehn Zehnerstäben und ein Hunderterwürfel aus zehn Hunderterplatten etc. besteht (Van de Walle, 2007). Fragen wie

„Warum heißt der Zehnerstab Zehnerstab? Warum heißt die Hunderterplatte Hunderterplatte? Warum heißt der Tausendwürfel Tausenderwürfel? Wer kann das mit dem Material begründen?“ (Freesemann, 2014, S. 103) können dafür hilfreich sein. Wichtig ist in einem weiteren Schritt, dass die Darstellung von Zahlen mithilfe des Dienes-Materials mit der Stellenwertschreibweise verknüpft wird (Freesemann, 2014). Dazu eignen sich insbesondere die Stellenwertkarten und die Stellenwerttafel (vgl. Abbildung 5).

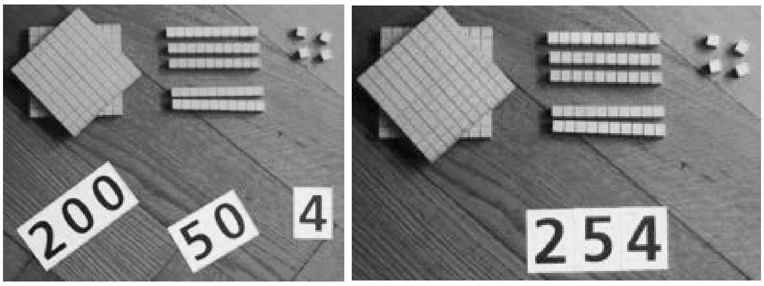


Abbildung 5: Darstellung der Zahl 254 mit dem Dienes-Material und den Stellenwertkarten

Mithilfe der Stellenwerttafel können zudem auch nicht standardisierte Notationsformen thematisiert werden (Freesemann, 2014; Scherer & Moser Opitz, 2010; Van de Walle, 2007). So kann die Zahl 254 sowohl als 2 Hunderter, 5 Zehner und 4 Einer notiert werden, es ist aber auch möglich 254 Einer oder 25 Zehner und 4 Einer oder 1 Hunderter, 14 Zehner und 14 Einer zu notieren (vgl. Abbildung 6). Werden solche Notationsformen von Schülerinnen und Schülern gewählt, können diese mit standardisierten Notationsformen verglichen und das Prinzip der fortgesetzten Bündelung an solchen Beispielen veranschaulicht und erarbeitet werden (Scherer & Moser Opitz, 2010). Auch ist an der Stellenwerttafel unter Zuhilfenahme des Dienes-Materials zu klären, was es heisst, wenn bei einem Stellenwert in der Tafel keine Ziffer eingetragen ist (z.B. 3 H, 4 E ist 304 und nicht 34; vgl. Scherer & Moser Opitz, 2010, S. 134). Zudem muss immer wieder die Verbindung zu den Zahlwörtern hergestellt (Van de Walle, 2007) und insbesondere im Deutschen auch thematisiert werden, dass bei den Zahlwörtern im Vergleich zur Ziffernschreibweise die Zehner-Einer-Inversion beachtet werden muss (Moser Opitz, 2013).

H (Hunderter)	**Z** (Zehner)	**E** (Einer)
2	5	4
		254
	25	4
1	14	14

Abbildung 6: Verschiedene Notationsformen für die Zahl 254 in der Stellenwerttafel

Ebenfalls explizit zu erarbeiten ist das Prinzip des Entbündelns (Scherer & Moser Opitz, 2010). Subtraktionsaufgaben sollen hierfür auch unter Zuhilfenahme von Arbeitsmitteln und Veranschaulichungen wie dem Dienes-Material, den Stellenwertkarten und/oder der Stellenwerttafel dargestellt und bearbeitet werden (Schmassmann & Moser Opitz, 2008b). Bei der Verwendung des Dienes-Materials muss beachtet werden, dass im Gegensatz zur Addition, bei der beide Summanden mit Material dargestellt werden, bei der Subtraktion nur der Minuend mit Material gelegt wird (Schmassmann & Moser Opitz, 2008b). Soll nun der Subtrahend vom Minuend subtrahiert werden, müssen bei Subtraktionen mit Überschreitungen Tausenderwürfel, Hunderterplatten bzw. Zehnerstäbe in kleinere Einheiten umgetauscht, also entbündelt werden (Scherer & Moser Opitz, 2010). Um das Umtauschen bzw. Entbündeln zu erarbeiten, werden Fragen wie „Wie viele Hunderterplatten bekomme ich für einen Tausenderwürfel? Wie viele Zehnerstäbe kriege ich für eine Hunderterplatte? Wie viele Einerwürfel für einen Zehnerstab?“ empfohlen (Schmassmann & Moser Opitz, 2008b, S. 92). Zudem ist es sinnvoll, systematische Veränderungen zu thematisieren, indem operativ strukturierte Aufgaben gestellt werden (z.B. 1000 – 100 = 900, 1000 – 10 = 990, 1000 – 1 = 999; vgl. Scherer & Moser Opitz, 2010, S. 143). Dabei kann mithilfe der Stellenwerttafel diskutiert und reflektiert werden, an welcher Stelle sich etwas ändert, wenn ein Einer, ein Zehner bzw. ein Hunderter abgezogen wird. Auch hier ist es wichtig, dass die Schülerinnen und Schüler dazu aufgefordert werden, ihre mathematischen Handlungen am Material zu versprachlichen (Scherer & Moser Opitz, 2010).

Ordinaler Zahlaspekt

Mit Blick auf den Aufbau eines tiefgehenden Verständnisses des dezimalen Zahlensystems ist zusätzlich zur bereits genannten, strukturorientierten Perspektive auch ein positionsorientiertes Verständnis von Zahlen zu erarbeiten (Freesemann, 2014). Für die Darstellung des ordinalen Zahlaspekts eignen sich insbesondere die Hunderterkette und der Zahlenstrahl, zudem können damit auch Grössenvorstellungen von Zahlen und die Orientierung im Zahlenraum gefördert werden (Freesemann, 2014). Bei der Hunderterkette handelt es sich um 100 Perlen, die auf einer Schnur aufgezogen und bei der jeweils immer 10 aufeinanderfolgende Perlen gleich eingefärbt sind, sodass sich z.B. jeweils 10 rote Perlen und 10 weisse Perlen abwechseln und diese als Zehnereinheit erfasst werden können (Höhtker & Selter, 1995). Für die Förderung der Orientierung im Zahlenraum machen Höhtker und Selter (1995) einen Vorschlag für eine Unterrichtsreihe, bei der ausgehend von der Hunderterkette der Zahlenstrahl erarbeitet wird. Auch hier sollte, wie auch bei anderen Veranschaulichungen und Arbeitsmitteln, zunächst die Hunderterkette mit den Schülerinnen und Schülern erarbeitet werden, indem der Aufbau und die Struktur der Hunderterkette erkundet werden (Höhtker & Selter, 1995). Freesemann (2014,

S. 107) schlägt dazu folgende Fragen vor: „Wie viele weiße und rote Perlen sind immer nebeneinander? Wie viele Perlen sind insgesamt auf der Kette? Wie können wir das schnell herausfinden? Wie viele Zehnerschritte zählen wir bis 100? Wie viele Zehner sind in einem Hunderter?“ Daran anschliessend sollen Orientierungsübungen an der Hunderterkette durchgeführt werden (Höhtker & Selter, 1995):

- Zahlen finden: Es wird ein Ort an der Hunderterkette vorgegeben, und es muss die passende Zahl gefunden werden,
- Orte finden: Es wird eine Zahl vorgegeben, und es muss der passende Ort an der Hunderterkette gefunden werden.

Wichtig ist dabei, mit den Schülerinnen und Schülern die Platzierung der Zahlen an der Hunderterkette zu thematisieren (so muss die Markierung für die Zahl 50 nach der 50. Perle gesetzt werden; Freesemann, 2014, S. 108). Im gemeinsamen Gespräch soll zudem herausgearbeitet werden, welche Stützpunkte für die Orientierung an der Hunderterkette sinnvoll zu nutzen sind (z.B. glatte Zehnerzahlen; Höhtker & Selter, 1995). Mit der Zeit sollen immer mehr Übungen auch an einer ikonischen Darstellung der Hunderterkette durchgeführt werden, um die Abstraktion vom konkreten Material zu fördern (Höhtker & Selter, 1995). Ebenfalls wird empfohlen, auch Übungen zum Finden von Nachbareinern, Nachbarzehnern und Nachbarhundertern von Zahlen durchzuführen (Freesemann, 2014). In einem nächsten Schritt wird ausgehend von der Hunderterkette dann der Zahlenstrahl erarbeitet. Dafür wird ein zunächst leerer Zahlenstrahl der Hunderterkette gegenübergestellt und an der Hunderterkette markierte Zahlen (z.B. glatte Zehnerzahlen) auf den leeren Zahlenstrahl übertragen (Höhtker & Selter, 1995). Danach sollen verschiedene Übungen für „Zahlen finden“ und „Orte finden“ am Zahlenstrahl durchgeführt und geeignete Stützpunkte für die Orientierung diskutiert werden (Freesemann, 2014; Höhtker & Selter, 1995). Für Schülerinnen und Schüler, die bereits im Tausenderraum arbeiten, sollte zudem auch dieser Zahlenraum entsprechend erarbeitet werden, indem zunächst zehn Hunderterketten aneinandergereiht werden und danach der Tausenderstrahl davon abgeleitet wird (Freesemann, 2014). Auch hier ist wieder zu diskutieren, welche Stützpunkte beim Tausenderstrahl hilfreiche Orientierungspunkte sind, und es können Aufgaben des Typs „Zahlen finden“ und „Orte finden“ durchgeführt werden.

Mathematisieren und Operationsverständnis

Wie in dieser Arbeit aufgezeigt werden konnte, haben rechenschwache Schülerinnen und Schüler Schwierigkeiten beim Mathematisieren und verfügen über ein eingeschränktes Operationsverständnis (vgl. Kap. 2.4.4). Die Begriffe „Mathematisieren“ und „Operationsverständnis“

können nicht losgelöst voneinander betrachtet werden, verweisen aber auf Unterschiedliches (Freesemann, 2014). Mathematisieren meint „reale Situationen in die Sprache der Mathematik übersetzen, mit Mitteln der Mathematik Lösungen bestimmen und das Ergebnis für die reale Situation interpretieren und umgekehrt“ und umfasst somit Kompetenzen, die für das Problemlösen und das Sachrechnen benötigt werden (Schmassmann & Moser Opitz, 2008a, S. 43).[11] Der Begriff des Operationsverständnisses hingegen verweist auf die vier Grundoperationen und das Verständnis für diese in folgendem Sinne:

> [Dabei] geht es darum, allen Kindern eine *Grundlegung des Operationsverständnisses* zu ermöglichen, d. h. Einsichten darüber zu gewinnen, was bspw. die Addition/die Subtraktion (als mathematische Idee) ausmacht, und wie man lernt, sie *flexibel* zu handhaben, was v. a. ein situationsbezogen *geschicktes* Rechnen unter Ausnutzung von Rechengesetzen und strukturellen Regelhaftigkeiten der jeweiligen Operation meint […]. (Krauthausen & Scherer, 2007, S. 24)

Das Mathematisieren kann somit als Teil des Operationsverständnisses angesehen werden. Für eine Förderung in der Primarschule wird insbesondere der Aufbau von Grundvorstellungen zu den vier Grundoperationen empfohlen (Häsel-Weide, 2016; Scherer & Moser Opitz, 2010). Dabei geht es darum, die vier Grundoperationen mit geeignetem Material bzw. mit Rechengeschichten zu veranschaulichen und so mit den Schülerinnen und Schülern Verknüpfungen zwischen verschiedenen Repräsentationsformen (konkrete Sachsituation, Modell/Bild, symbolische Darstellung) zu erarbeiten (Scherer & Moser Opitz, 2010). Hierbei wird als wichtig erachtet, dass die verschiedenen Repräsentationsformen miteinander verglichen, mathematische Handlungen bzw. Operationen versprachlicht und reflektiert werden (Scherer & Moser Opitz, 2010; Wartha & Schulz, 2011). Das Hinzufügen und Wegnehmen bei der Addition und Subtraktion kann z.B. mithilfe von Plättchen gut dargestellt werden, gleichzeitig muss aber auf eine strukturierte Darstellungsweise geachtet werden, sodass nicht in Einerschritten dazu- oder weggezählt wird (Häsel-Weide, 2016). Bei der Multiplikation müssen zwei Modellvorstellungen berücksichtigt werden. So eignen sich für die Darstellung der Multiplikation im Sinne des räumlich simultanen Modells das Punktefeld, aber auch Bilder mit multiplikativer Struktur aus dem Alltag (z.B. Eierkartons mit 6 oder 8 Eiern auf 2 Reihen verteilt; Scherer & Moser Opitz, 2010). Ebenso sollten die Multiplikationen auch als zeitlich-sukzessives Modell veranschaulicht werden, wobei sich dafür Rechengeschichten eignen (z.B. „Vier mal zwei Flaschen in den Keller bringen“; Krauthausen & Scherer, 2007, S. 28). Auch bei der Division müssen zwei unter-

[11] Für eine ausführliche Darstellung von Modellierungsprozessen beim Mathematisieren siehe Leiss (2010) und Blum und Leiss (2005).

schiedliche Modellvorstellungen beachtet werden (Krauthausen & Scherer, 2007). Einerseits geht es bei der Division ums Aufteilen (z.B. „In einer Turnhalle sind 20 Kinder. […] Wie viele Vierer-Gruppen können gebildet werden?“ → Es wird die Anzahl Teilmengen gesucht; Krauthausen & Scherer, 2007, S. 30), andererseits kann die Division auch als Verteilen dargestellt werden (z.B. „In einer Turnhalle sind 20 Kinder. Es sollen vier gleich grosse Gruppen gebildet werden. Wie viele Kinder sind in einer Gruppe?“ → Es wird die Elementeanzahl der Teilmengen gesucht; Krauthausen & Scherer, 2007, S. 30).

Insbesondere beim Zeichnen (aber auch beim Versprachlichen) von Sachsituationen muss jedoch auf die Mehrdeutigkeit entsprechender Darstellungen geachtet werden (Scherer & Moser Opitz, 2010). So kann die Situation „Fünf Vögel sitzen auf einem Baum, drei sind in der Luft“ sowohl als Addition (5 + 3 = 8) als auch als Subtraktion (8 – 3 = 5 bzw. 8 – 5 = 3) als auch als Ergänzungsrechnung (5 + ? = 8 bzw. 3 + ? = 8) notiert oder in Worte gefasst werden (Scherer & Moser Opitz, 2010). Auch bei der Darstellung der Subtraktion kann der dynamische Prozess des Wegnehmens unterschiedlich veranschaulicht und verstanden werden (Kinder argumentieren z.B., dass leer getrunkene Gläser immer noch vorhanden sind und daher nichts weggenommen wurde; vgl. Scherer & Moser Opitz, 2010, S. 81). Es wird deshalb allgemein gefordert, „[…] dass Lehrpersonen die Vorstellungsbilder der Schülerinnen und Schüler erfragen und sie wenn nötig beim Aufbau von alternativen Repräsentationen unterstützen und begleiten“ (Scherer & Moser Opitz, 2010, S. 81). Die Mehrdeutigkeit von Darstellungen soll dabei im Unterricht explizit genutzt werden, indem zur gleichen Darstellung verschiedene zutreffende Aufgaben gesucht und individuelle Sichtweisen gemeinsam reflektiert werden (Scherer & Moser Opitz, 2010).

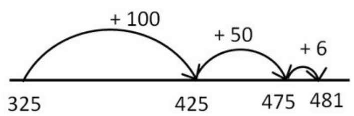
Diesbezüglich ist auch zu beachten, dass insbesondere bei rechenschwachen Schülerinnen und Schülern Handeln nicht automatisch zu Einsicht führt, sondern dass dieser Prozess aktiv unterstützt und methodisch durchdacht begleitet werden muss, damit die Kinder die hinter den Operationen stehenden Handlungen verinnerlichen können (Häsel-Weide, 2016; Koch, 2008; Scherer & Moser Opitz, 2010; Wartha & Schulz, 2011). Um die Kinder für die konkreten (Mengen-) Veränderungen zu sensibilisieren, die die Handlungen beinhalten, kann das Vierphasenmodell zum Aufbau von Grundvorstellungen von Wartha und Schulz (2011, S. 11) hilfreich sein. Demnach soll zunächst mit geeignetem Material gehandelt, dabei aber die Handlungen und die mathematischen Symbole versprachlicht werden. In einem zweiten Schritt sollen die Handlungen am Material nicht mehr selbst, sondern durch jemand anderes ausgeführt werden, sodass die Handlungen beobachtet und kontrolliert werden können. In einem dritten Schritt soll das

Kind die Handlungen ohne Sicht auf das Material beschreiben, sodass die Handlungsprozesse innerlich vorgestellt werden müssen. Erst in einem vierten Schritt soll auf rein symbolischer Ebene gearbeitet, geübt und automatisiert werden, wobei die Handlungen bei Bedarf wiederum innerlich vorgestellt werden können (Wartha & Schulz, 2011). Mit solchen Übungen können Mathematisierungskompetenzen gefördert werden, und es wird dazu beigetragen, dass langfristig ein tragfähiges Operationsverständnis aufgebaut werden kann.

Halbschriftliche Rechenstrategien

Wie im Kapitel 2.4.3 ausgeführt wurde, weisen rechenschwache Kinder auch einen eingeschränkten Strategiengebrauch auf. Aus mathematikdidaktischer Sicht ist es daher wichtig, rechenschwache Kinder mit Strategien vertraut zu machen, die ihnen helfen können, auch mehrstellige Rechenoperationen korrekt durchzuführen. Hierfür eignen sich insbesondere halbschriftliche Verfahren, da diese das Arbeitsgedächtnis entlasten, weil Zwischenschritte und Zwischenresultate notiert werden können (z.B. Scherer & Moser Opitz, 2010). An dieser Stelle sollen vier halbschriftliche Hauptstrategien vorgestellt werden, die für die Primarschule besonders relevant sind und von Wittmann und Müller (1992) systematisiert wurden (vgl. Tabelle 2). Dabei ist zu beachten, dass im Einzelfall auch mehr oder weniger abweichende Notationsformen gewählt werden können (Krauthausen & Scherer, 2007). Die Strategie *Stellenwerte extra* wird sehr häufig eingesetzt und wird manchmal auch dann von Kindern verwendet, wenn sie nicht explizit in der Schule eingeführt wurde (Selter, 2000). Bei dieser Strategie werden beide Zahlen, die addiert oder subtrahiert werden müssen, jeweils in die einzelnen Einer, Zehner, Hunderter usw. zerlegt und dann die Einer, Zehner und Hunderter separat addiert oder subtrahiert (Wittmann & Müller, 1992; vgl. Tabelle 2). Dabei ist jedoch für die Subtraktion zu berücksichtigen, dass diese Strategie speziell bei Aufgaben mit Überschreitungen auch ihre Tücken hat (Scherer & Moser Opitz, 2010; vgl. Tabelle 2). So müssen einerseits die Teilergebnisse für die verschiedenen Stellenwerte jeweils addiert und nicht subtrahiert werden (Scherer & Moser Opitz, 2010). Andererseits besteht bei Aufgaben mit Überschreitungen die Problematik, dass negative Zwischenergebnisse resultieren, was Kinder dazu verleiten kann, den Minuend und den Subtrahend zu vertauschen, sodass sie trotz fehlerhaftem Vorgehen eine korrekte Lösung erzielen können (z.B. $53 - 28 = ?$, $50 - 20 = 30$, $8 - 3 = 5$, $30 - 5 = 25$; vgl. Scherer & Moser Opitz, 2010, S. 151). Für schwache Lernende ist die Strategie *Schrittweise*, bei der von der unveränderten Startzahl die Hunderter, die Zehner und die Einer separat addiert oder subtrahiert werden, in vielen Fällen das geeignetere Vorgehen, weil weniger Teilschritte notwendig sind und eine Veranschaulichung der einzelnen Rechenschritte am Rechenstrich möglich ist (Scherer & Moser Opitz, 2010; vgl. Tabelle 2). Anspruchsvoller sind die Strategien Vereinfachen und Hilfsaufgabe (Scherer & Moser Opitz, 2010). Beim *Vereinfachen* werden die Summanden auf Basis der Konstanz der Summe so vereinfacht, dass eine Rechnung entsteht, die leichter gelöst werden kann (Wittmann & Müller, 1992; vgl. Tabelle 2). Bei der Strategie *Hilfsaufgabe* hingegen wird eine benachbarte Rechnung gesucht, die leichter zu berechnen ist, und von dieser wird dann das Ergebnis für die eigentliche Aufgabe abgeleitet (Wittmann & Müller, 1992; vgl. Tabelle 2). Da für die beiden letztgenannten halbschriftlichen Rechenstrategien Zahlbeziehungen genutzt werden müssen, sind auch diese Strategien für rechenschwache Kinder ohne spezifische Erarbeitung eher schwer zu nutzen (Scherer & Moser Opitz, 2010).

Tabelle 2: Übersicht zu häufig verwendeten halbschriftlichen Rechenverfahren (vgl. Wittmann & Müller, 1992; Scherer & Moser Opitz, 2010)

Strategie	**Beispiel**	**Bemerkungen**
Stellenwerte extra	325 + 156 = ? 300 + 100 = 400 20 + 50 = 70 5 + 6 = 11 400 + 70 + 11 = 481	Problematisch bei Subtraktionen mit Überschreitungen, da bei diesen mit negativen Zahlen gerechnet werden muss: 53 – 28 = ? 50 – 20 = 30 3 – 8 = –5 30 – 5 = 25
Schrittweise	325 + 156 = ? 325 + 100 = 425 425 + 50 = 475 475 + 6 = 481	Dieses Verfahren kann auf dem Rechenstrich gut veranschaulicht werden: 
Vereinfachen	325 + 156 = ? 330 + 151 = 481	Zahlbeziehungen müssen genutzt werden (anspruchsvoll)
Hilfsaufgaben	325 + 156 = ? 330 + 156 = 486 → 325 + 156 = 481	Zahlbeziehungen müssen genutzt werden (anspruchsvoll)

Ein wichtiger Aspekt bei der Erarbeitung halbschriftlicher Rechenstrategien ist der Einsatz von geeigneten Veranschaulichungen und Hilfsmitteln (z.B. Dienes-Material, Rechenstrich; Scherer & Moser Opitz, 2010). So kann unter Zuhilfenahme des Dienes-Materials bei der Subtraktion thematisiert werden, dass der Minuend und der Subtrahend nicht vertauscht werden dürfen und was es bedeutet, wenn der Einer-Subtrahend kleiner ist als der Einer-Minuend. Die oben bereits erwähnte Rechnung (53 – 28 = ?) könnte folgendermassen bearbeitet werden: „Wir legen zuerst die Ausgangszahl mit Zehnerstäben und Einerwürfeln. Wir haben also fünf Zehnerstäbe und drei Einerwürfel, nun müssen wir acht Einerwürfel wegnehmen. D.h., wir müssen zuerst einen der fünf Zehnerstäbe nehmen und diesen zu zehn Einerwürfeln umtauschen. Nun können wir von den dreizehn Einerwürfeln acht wegnehmen. Welche Zahl erhalten wir als Zwischenergebnis? Und nun nehmen wir noch zwei Zehner weg. Welche Zahl erhalten wir jetzt?“

Auch bei der Strategie *Schrittweise* kann es insbesondere für rechenschwache Schülerinnen und Schüler notwendig sein, Arbeitsmittel einzusetzen. So kann die Hunderterkette mit gefärbten Fünfer- und Zehnerkugeln dabei helfen, zu erkennen, was es heisst, wenn zu gemischten Zehnerzahlen beliebig viele Zehner addiert werden (Scherer & Moser Opitz, 2010). In einem zweiten Schritt kann dann unter Zuhilfenahme des Rechenstrichs gezeigt werden, dass mehrere Zehnersprünge auch zu einem Sprung zusammengefasst werden können (Scherer & Moser Opitz, 2010).

Wichtig ist zudem, dass in Phasen der Reflexion im Klassenverband immer wieder verschiedene Rechenstrategien und Notationsformen benannt, miteinander verglichen und gemeinsam besprochen werden (Scherer & Moser Opitz, 2010). Das Verbalisieren eigener Rechen- und Denkwege wird auch bei rechenschwachen Schülerinnen und Schülern als wichtiger Bestandteil für den Aufbau von konzeptuellem Verständnis angesehen (Werner, 2009). Allerdings weisen Schmassmann und Moser Opitz (2008b, S. 44) darauf hin, dass nicht „[…] alle Schülerinnen und Schüler alle Wege nachvollziehen müssen bzw. in der Lage sein sollen, alle Wege selbst anzuwenden“. Für rechenschwache Kinder ist es wichtig, dass sie zunächst einen für sie passenden Weg finden und lernen, diesen anzuwenden, bevor auch andere Wege ausprobiert werden können (Schmassmann & Moser Opitz, 2008b).

Zusammenfassend lässt sich somit sagen, dass aus mathematikdidaktischer Sicht unter Berücksichtigung rechenschwacher Schülerinnen und Schüler eine Mathematikförderung auf zentrale mathematische Inhalte fokussieren, einen systematischen Kompetenzaufbau verfolgen, produktive Übungsformate berücksichtigen und geeignete Veranschaulichungen und Arbeitsmittel verwenden sollte. Mit Blick auf spezifische Inhalte sind zudem geeignete Massnahmen zu treffen, die eine Ablösung vom zählenden Rechnen unterstützen und ein umfassendes Verständnis des dezimalen Stellenwertsystems und des kardinalen Zahlaspekts ermöglichen. Ebenso ist der ordinalen Zahlaspekt mit geeigneten Massnahmen zu thematisieren, und es müssen Mathematisierungskompetenzen sowie das Operationsverständnis gefördert und geeignete halbschriftliche Rechenstrategien erarbeiten werden. Die in diesem Kapitel dargestellten mathematikdidaktischen Überlegungen werden für die konkrete Gestaltung der in dieser Arbeit zu untersuchenden Mathematikförderung daher berücksichtigt. Eine ausführliche Beschreibung, wie diese Aspekte in der vorliegenden Interventionsstudie umgesetzt werden, findet sich im Kapitel 5.3.1.

3.4 Bedeutung von professionellem Wissen und professionellen Kompetenzen der Lehrpersonen

In der vorliegenden Arbeit wurden bisher empirische Erkenntnisse sowie didaktische und mathematikdidaktische Überlegungen zur unterrichtsintegrierten Förderung von rechenschwachen Kindern in regulären Schulklassen vorgestellt. Dabei wurde bereits an verschiedenen Stellen darauf hingewiesen, dass Lehrpersonen dafür spezifische Kompetenzen und Wissensbestände benötigen. In diesem Kapitel werden professionelles Wissen und professionelle Kompetenzen der Lehrpersonen aus theoretischer Sicht näher erläutert, und es wird auf die Bedeutung dieser Aspekte für die Lern- und Leistungsentwicklung von Schülerinnen und Schülern eingegangen. Abschliessend werden Professionalisierungsmassnahmen vorgestellt, mit denen das Wissen und die Kompetenzen der Lehrpersonen gesteigert werden können.

Wenn es um das professionelle Wissen der Lehrpersonen geht, wird meistens auf die Konzeption von Shulman (1986; 1987) verwiesen, in der zwischen Fachwissen, fachdidaktischem Wissen, curricularem Wissen und allgemeinem pädagogischem Wissen unterschieden wird. Mit dem *Fachwissen* ist das Wissen einer Lehrperson gemeint, das sich aus dem Faktenwissen eines Faches sowie dessen Konzepten und Prinzipien, auch mit Bezug auf deren innere Zusammenhänge und deren Relevanz in Theorie und Praxis, zusammensetzt (Shulman, 1986). Das *fachdidaktische Wissen* verweist darauf, dass sich eine Lehrperson darüber hinaus damit auskennen muss, welche Darstellungsmöglichkeiten für den Unterricht besonders geeignet sind und welche inhaltlichen Schwierigkeiten bei Schülerinnen und Schülern beim Lernen eines bestimmten Lerngegenstandes auftreten können (Shulman, 1986). Mit *curricularem Fachwissen* ist das Wissen einer Lehrperson zum Lehrplan und zum Curriculum gemeint, das jedem Unterricht zugrunde liegt, diesen organisiert und die zu lehrenden Inhalte bestimmt (Shulman, 1986). Dies umfasst auch das Wissen über geeignete Materialien und Lehrmittel sowie über Verbindungen zwischen verschiedenen Lerninhalten, die zu unterschiedlichen Zeitpunkten oder in unterschiedlichen Fächern gelehrt werden (Shulman, 1986). Das *allgemeine pädagogische Wissen* verweist auf notwendiges Wissen zur Klassenführung, zur Organisation von Unterricht und zu Lernprozessen im Allgemeinen (Shulman, 1986). In neueren Untersuchungen zum Wissen und zu Kompetenzen der Lehrpersonen wie z.B. in COACTIV wird die Konzeption von Shulman (1986; 1987) aufgegriffen, jedoch noch durch weitere Merkmale der Lehrpersonen, die als relevant für eine erfolgreiche Lehrtätigkeit angesehen werden, ergänzt. Gemäss COACTIV setzen sich die professionellen Kompetenzen der Lehrpersonen demnach aus Professionswissen, Werthaltungen und Überzeugungen, motivationalen Orientierungen und Selbstregulation zusammen (Baumert & Kunter, 2011).

Zur Bedeutung professionellen Wissens der Lehrpersonen für die Lernerfolge der Schülerinnen und Schüler liegen inzwischen einige Studien vor. So konnten Hill, Rowan und Loewenberg Ball (2005) für den Mathematikunterricht zeigen, dass fachdidaktisches Wissen unter Kontrolle diverser Lernvoraussetzungen aufseiten der Schülerinnen und Schüler sich positiv auf den Lernerfolg ebendieser in der ersten und der dritten Klasse auswirkt. Beim fachdidaktischen Wissen handelte es sich dabei um den stärksten lehrpersonenbezogenen Prädiktor, der untersucht wurde (Hill et al., 2005). Auch in COACTIV konnte ein positiver Einfluss des mathematikdidaktischen Wissens auf die Leistungsentwicklung der Schülerinnen und Schüler belegt werden, wobei 39% der Leistungsvarianz zwischen den Klassen durch das mathematikdidaktische Wissen erklärt werden konnten (Kunter et al., 2011). Für das Fachwissen liegen weniger eindeutige Befunde vor. So zeigte sich in COACTIV nur ein indirekter Einfluss des mathematischen Fachwissens auf die Leistungsentwicklung in Mathematik, sodass sich ein hohes mathematisches Fachwissen positiv auf das mathematikdidaktische Wissen auswirkt und dieses wiederum die Leistungen der Schülerinnen und Schüler positiv beeinflusst (Kunter et al., 2011). Eine weitere Untersuchung belegt zwar auch einen leicht positiven Einfluss des mathematische Fachwissens auf die Leistungsentwicklung der Schülerinnen und Schüler, allerdings wurde das Fachwissen dort nur indirekt über vorhandene Fachabschlüsse der Lehrpersonen erfasst (Goldhaber & Brewer, 2000). Für das allgemeine pädagogische Wissen liegt zudem eine Studie vor, die zeigt, dass dieses Wissen in einem positiven Zusammenhang mit Qualitätsaspekten von Unterricht (Klassenführung, konstruktive Lernunterstüztung), beurteilt durch die Schülerinnen und Schüler, steht (Voss, Kunter, Seiz, Hoehne & Baumert, 2014). Da das professionelle Wissen und die Kompetenzen der Lehrpersonen bedeutsam für die Leistungsentwicklung der Schülerinnen und Schüler sind, stellt sich die Frage, inwiefern diese Aspekte z.B. im Rahmen von Fort- und Weiterbildungen der Lehrpersonen gefördert werden können.

Die empirische Befundlage zur Wirksamkeit institutionalisierter Fort- und Weiterbildungsmassnahmen weist auf das positive Potenzial solcher Angebote zur Steigerung professionellen Wissens und professioneller Kompetenzen der Lehrpersonen ebenso wie zur Steigerung der Leistungen der Schülerinnen und Schüler hin (Lipowsky, 2011). Auf didaktischer Ebene ist dabei zentral, dass Professionalisierungsmassnahmen einen eindeutigen curricularen Bezug haben, domänenspezifisch sind und auf evidenzbasierten Aspekten eines erfolgreichen Unterrichts aufbauen (Lipowsky, 2011). Weiter ist wichtig, dass der Fokus auf die Lern- und Verstehensprozesse der Schülerinnen und Schüler gerichtet ist und reflektiert wird, wie diese mit dem unterrichtlichen Handeln der Lehrperson in Verbindung stehen (Lipowsky, 2011). Für den Mathematikunterricht konnten Carpenter, Fennema, Peterson, Chiang und Loef (1989) in einer älteren Untersuchung zeigen, dass eine vierwöchige Weiterbildung wirksam war, die auf eine Verbesserung des fachdidaktischen und des diagnostischen Wissens der Lehrpersonen setzte, indem Lern- und Verstehensprozesse von Schülerinnen und Schülern (u.a. in Form von Videoaufnahmen) als Ausgangspunkt für die Reflexion des Handelns der Lehrperson in den Mittelpunkt gestellt wurden. Sowohl auf Ebene der Lehrpersonen als auch auf Ebene der Schülerinnen und Schüler konnten für die Experimentalgruppe positive Effekte belegt werden (Carpenter et al., 1989). Auch eine neuere Untersuchung zum Mathematikunterricht konnte aufzeigen, dass ein Einsatz von Videosequenzen mit Fokus auf die kognitive Aktivierung von Schülerinnen und Schülern und deren Optimierungsmöglichkeiten zu einer veränderten Sicht auf den Unterricht geführt hat, wobei jedoch nicht untersucht wurde, ob dadurch die Unterrichtsgestaltung auch tatsächlich positiv beeinflusst wurde, ebenso fehlte eine Kontrollgruppe (Krammer, Lipowsky, Pauli, Schnetzler & Reusser, 2012). Eine Fokussierung auf eine inhaltsspezifische Planung und Reflexion von Lehr-Lern-Prozessen ist auch der Kern des Ansatzes des fachspezifischen Unterrichtscoaching (Staub, 2014). Bei diesem wird jedoch noch viel stärker auf die Reflexion des eigenen, konkreten Unterrichts in einem konstanten Tandem geachtet wird, indem der Unterricht und die dabei intendierten Lehr-Lern-Prozesse mit einem Coach oder einem Peer aus dem Kollegium gemeinsam geplant, gemeinsam durchgeführt, gemeinsam nachbearbeitet und reflektiert werden (Staub, 2014).

Zusammenfassend kann somit festgehalten werden, dass Lehrpersonen u.a. über fachliches, fachdidaktisches, curriculares und allgemeinpädagogisches Wissen verfügen müssen, um den Unterricht erfolgreich gestalten zu können. Mit Blick auf die mathematischen Leistungen der Schülerinnen und Schüler muss zudem das mathematikdidaktische Wissen als zentral erachtet werden. Aus dem Forschungsdiskurs zu Professionalisierungsmassnahmen ist zudem bekannt, dass Lehrpersonen im Rahmen von Fortbildungen in ihrem Wissen und ihren Kompetenzen gestärkt werden können. Wichtig ist dabei, dass eine inhaltliche Fokussierung auf die Gestaltung des Unterrichts erfolgt sowie Gelegenheiten zum Reflektieren geboten werden und dass sowohl das unterrichtliche Handeln der Lehrperson als auch die Lern- und Verstehensprozesse der Schülerinnen und Schüler thematisiert werden.

4. Zwischenfazit und Schlussfolgerungen

In der vorliegenden Arbeit wurde bisher eine theoretische und empirische Eingrenzung des Phänomens Rechenschwäche vorgenommen, und es wurden Erkenntnisse zur Förderung von Schülerinnen und Schülern mit unterdurchschnittlichen bis stark unterdurchschnittlichen Mathematikleistungen präsentiert. Die zentralsten Überlegungen werden hier zuerst nochmals zusammenfassend dargestellt sowie offene Fragen im Bereich der Erforschung der Wirksamkeit von mathematischen Förderansätzen thematisiert. Im Anschluss daran wird aufgezeigt, wie basierend auf den bisherigen Erkenntnissen dieser Arbeit eine verstehensorientierte Mathematikförderung für die dritte Klasse, die im regulären Unterricht der Primarschule integriert wird, gestaltet werden soll. Zudem wird aufgezeigt, wie die Förderung im Rahmen des Angebots-Nutzungs-Modells (Fend, 1981; Helmke & Weinert, 1997) eingeordnet werden kann.

4.1 Zwischenfazit und offene Fragen

4.1.1 Rechenschwäche als Phänomen

Wie im Kapitel 2.1.3 ausgeführt, weisen als rechenschwach bezeichnete Schülerinnen und Schüler unterdurchschnittliche bis stark unterdurchschnittliche Mathematikleistungen im Vergleich zu ihren Mitschülerinnen und Mitschülern auf. Ein enges Begriffsverständnis im Sinne der ICD-10-Definition von Rechenstörung wird für die vorliegende Arbeit auf Basis der im Kapitel 2.1 geäusserten Überlegungen und Kritikpunkte (vgl. Ehlert et al., 2012; Moser Opitz & Ramseier, 2012) hingegen nicht als sinnvoll erachtet.

Es wird davon ausgegangen, dass verschiedene Faktoren an der Entstehung einer Rechenschwäche beteiligt sein können (vgl. Kap. 2.3 und 3.4). Zu den zentralen Einflussfaktoren, die zu einer schwachen Mathematikleistung beitragen, gehören neben domänenübergreifenden und domänenspezifischen kognitiven Faktoren wie dem Arbeitsgedächtnis und dem Zahlensinn die Intelligenz, das mathematische Vorwissen, sprachliche Faktoren, das Geschlecht und damit in Verbindung stehende soziokulturelle Faktoren, der sozioökonomische Status, motivationale und emotionale Faktoren, aber auch die Qualität des Mathematikunterrichts sowie professionelle Kompetenzen und professionelles Wissen der Lehrpersonen (z.B. Brühwiler et al., 2017; de Smedt et al., 2013; Else-Quest et al., 2010; Geary et al., 2009; Hill et al., 2005; Paetsch et al., 2016; W. Schneider et al., 2013). Wie gross der Einfluss der einzelnen Faktoren auf die Entwicklung einer Rechenschwäche ist und wie sich die einzelnen Faktoren gegenseitig beeinflussen, ist bis heute jedoch noch weitgehend ungeklärt. Als Konsens kann allerdings angenommen werden, dass einer Rechenschwäche eine multikausale Verursachung zugrunde liegt und somit individuelle, soziale, unterrichtliche und gesellschaftliche Faktoren an der Entstehung einer Rechenschwäche beteiligt sind (vgl. z.B. Moser Opitz, 2013).

Im Kapitel 2.4 wurde aufgezeigt, dass es gewisse mathematische Inhalte und Kompetenzen gibt, bei denen rechenschwache Schülerinnen und Schüler grössere Schwierigkeiten haben als Schülerinnen und Schüler ohne Rechenschwäche. So können sich erste Schwierigkeiten bereits vor dem Eintritt in die Schule im Bereich der mathematischen Vorläuferfertigkeiten zeigen (z.B. Zählkompetenz, Verständnis für die Mächtigkeit von Zahlen; vgl. Krajewski & Schneider, 2009). Mit Blick auf die Zählkompetenz bereitet rechenschwachen Schülerinnen und Schülern das Zählen in Schritten grösser als eins und das Rückwärtszählen zudem auch in höheren Klassenstufen noch Mühe (z.B. Moser Opitz, 2013). Ebenso haben rechenschwache Schülerinnen und Schüler auch Schwierigkeiten beim Abruf von arithmetischem Faktenwissen und setzen häufiger zählende Rechenstrategien ein, sodass ein wenig effizienter Strategiengebrauch angenommen werden muss (z.B. Ostad, 1997). Sowohl beim Faktenabruf als auch beim zählenden Rechnen ergeben sich zudem höhere Fehlerquoten (Geary, Hoard & Bailey, 2012). Ebenso zeigen Studien, dass Schülerinnen und Schüler mit Rechenschwäche über ein eingeschränktes Operationsverständnis verfügen und Schwierigkeiten beim Mathematisieren bekunden (z.B. Andersson, 2008a). Zudem haben rechenschwache Schülerinnen und Schüler auch weniger konzeptuelles Verständnis des dezimalen Stellenwertsystems (z.B. Jordan & Hanich, 2003). Bei all diesen Aspekten handelt es sich um mathematische Inhalte bzw. Kompetenzen, die für die weitere mathematische Entwicklung als wesentlich gelten und somit als „Stolpersteine“ mathematischer Lernprozesse erachtet werden können (Moser Opitz, 2013, S. 143). Inhaltlich sollte eine Förderung rechenschwacher Schülerinnen und Schüler daher an diesen empirischen Erkenntnissen ansetzen und zusätzlich auch Modelle der mathematischen Entwicklung und Verarbeitung berücksichtigen (vgl. z.B. W. Schneider et al., 2013), die insbesondere für die frühe mathematische Entwicklung einige wichtige Hinweise bieten (z.B. Bedeutsamkeit des Erlernens der Zahlwortreihe, des Aufbaus eines mentalen Zahlenstrahls und des Teile-Ganzes-Verständnisses; vgl. Kap. 2.2).

4.1.2 Mathematische Förderung bei Rechenschwäche

Wie in dieser Arbeit aufgezeigt wurde, sind inzwischen auch einige Erkenntnisse zur wirksamen Förderung bei Rechenschwäche bekannt. So weist eine Metastudie deutschsprachiger Untersuchungen darauf hin, dass Einzelförderung wirksamer ist als Kleingruppenförderung oder klassenintegrierte Förderung (Ise et al., 2012). Diesbezüglich muss jedoch beachtet werden, dass gemeinsame Lernsituationen einerseits für mathematische Lernprozesse als wichtig erachtet werden (z.B. Freudenthal, 1974), andererseits auch gemeinsamkeitsstiftende Effekte haben, was für die Ziele eines integrativen bzw. inklusiven Schulsystems bedeutsam ist (z.B. Wocken, 1998). Zudem konnten Wiener und Tardif (2004) zeigen, dass eine Förderung, die im regulären Klassenraum anstatt in sogenannten „ressource rooms“ stattfindet, sich positiv auf die soziale Akzeptanz und das mathematische Selbstkonzept auswirkt. Weiter haben akademische Gruppenaktivitäten einen positiven Einfluss auf schulische Leistungen und soziale Interaktionen von Kindern mit Lern- bzw. Verhaltensschwierigkeiten (Garrote et al., 2017). Damit gemeinsame Lernsituationen für individuelle Lernprozesse produktiv genutzt werden können, muss im Mathematikunterricht somit ein didaktisches Setting geschaffen werden, das eine Auseinandersetzung an einem gemeinsamen Gegenstand ermöglicht (Feuser, 2013).

Stark strukturierte Instruktionsansätze haben sich für die Förderung rechenschwacher Schülerinnen und Schüler in Metastudien als äusserst wirksam erwiesen (Gersten et al., 2009; Kroesbergen & van Luit, 2003). Wie im Kapitel 3.2.1 aufgezeigt wurde, werden solche stark strukturierten und lehrpersonengesteuerten Ansätze jedoch für heterogene Lerngruppen auch als problematisch angesehen, da individuelle Lern- und Verstehensprozesse in diesen nur in geringem Masse berücksichtigt werden (z.B. Montague, 2011). Zudem wird kritisiert, dass sich diese Ansätze für den Aufbau konzeptuellen Verständnisses weniger gut eignen, da ein tiefergehender Verständnisaufbau und ein flexibles Anwenden von Gelerntem vernachlässigt würden (z.B. Hasselhorn & Gold, 2017; Pfister et al., 2015). Unklar bleibt ausserdem, ob diese Förderansätze, die stark auf „drill and practice“ setzen (Montague, 2011, S. 295), nur sehr isoliert einzelne Rechenfertigkeiten fördern, und ob sie auch längerfristig wirksam sind (vgl. Kap. 3.2.1). Als wichtig können daher Strukturierungsmassnahmen einer adaptive Lernbegleitung erachtet werden, die sowohl eine individuelle Lernbegleitung als auch ein hohes Ausmass an Strukturierung von Lernprozessen ermöglichen und stärker auf einem verstehensorientierten Verständnis von Lernen aufbauen (vgl. Kap. 3.2.1). Im Zentrum einer adaptiven Lernbegleitung steht dabei die Frage nach der optimalen Passung zwischen den Hilfestellungen der Lehrperson und den individuellen Lernvoraussetzungen der Schülerinnen und Schüler (Parsons et al., 2018), was unter Berücksichtigung schwacher Lernenden als zentral beurteilt werden kann. Gemeinsam ist beiden Ansätzen, dass das Geben gezielter Hinweise, das Lernen am Modell und die verbale Begleitung von Lernprozessen als Massnahmen zur Strukturierung von Lernprozessen zum Einsatz kommen, wobei diese Strategien bei der adaptiven Lernbegleitung stärker am Individuum ansetzen (vgl. Kap. 3.2.1). Als bedeutsame Facetten für eine adaptive Lernbegleitung im Mathematikunterricht gelten folgende Aspekte: eine kognitive Aktivierung durch herausfordernde Aufgaben und Fragen, ein Anregen von Diskursen mit dem Ziel der Reflexion, ein produktiver

Umgang mit Fehlern als Ansatzpunkt für individuelle Unterstützung, Zielorientierung durch das Geben konkreter Hinweise sowie ein produktiver Einsatz von Arbeitsmitteln für den Verständnisaufbau (vgl. Pfister et al., 2015).

Inzwischen gibt es zusätzlich zu oben genannten Metastudien auch einzelne Interventionsstudien, die zeigen, dass verstehensorientierte Förderansätze für rechenschwache Schülerinnen und Schüler wirksam sind. Demnach ist ein entwicklungsorientiertes Training zentraler Vorläuferfertigkeiten (z.B. Teile-Ganzes-Verständnis) in Kleingruppen zu Beginn der Schulzeit bei rechenschwachen Kindern wirksam (vgl. Kap. 3.1.2). Auch eine verstehensorientierte Förderung arithmetischer Kompetenzen rechenschwacher Schülerinnen und Schüler in Kleingruppen in der Unterstufe kann auf Basis der in Kapitel 3.1.3 zitierten Interventionsstudien als wirksam angesehen werden, wobei positivere Effekte für längere Förderzeiträume anzunehmen sind (vgl. Ise et al., 2012). Für die Mittelstufe liegen zudem Interventionsstudien für rechenschwache Schülerinnen und Schüler vor, die auch bei (teilweise) klassenintegrierten, verstehensorientierten Förderansätzen deren Wirksamkeit belegen konnten (vgl. Kap. 3.1.4). Weitere empirische Hinweise für die Wirksamkeit von verstehensorientierten Massnahmen im Rahmen mathematischer Förderung finden sich zudem in Metastudien. Demnach sind Förderansätze für rechenschwache Schülerinnen und Schüler erfolgreich, in denen Schülerinnen und Schüler zum Verbalisieren angeregt werden, in denen eine produktive Übungspraxis eingesetzt wird, Schülerinnen und Schüler Feedback erhalten, Visualisierungen gezielt eingesetzt und Übungsaufgaben speziell ausgewählt und sequenziert werden (Gersten et al., 2009). In diesem Zusammenhang sind auch mathematikdidaktische Überlegungen zu beachten. Wie im Kapitel 3.3 an verschiedenen Stellen thematisiert wurde, werden eine Versprachlichung mathematischer Handlungen und darauf aufbauende Reflexionsprozesse aus mathematikdidaktischer Sicht als wichtige Massnahme für einen verstehensorientierten Mathematikunterricht insbesondere auch unter Berücksichtigung von Schülerinnen und Schülern mit Rechenschwäche angesehen (z.B. Scherer & Moser Opitz, 2010). Produktives Üben und der Einsatz geeigneter Veranschaulichungen und Hilfsmittel werden im mathematikdidaktischen Diskurs ebenfalls als wichtig erachtet, um den Aufbau von Einsicht und Verständnis zu ermöglichen (vgl. Kap. 3.3.1).

Wie in dieser Arbeit verschiedentlich aufgezeigt wurde, benötigen Lehrpersonen für einen erfolgreichen Mathematikunterricht im Allgemeinen sowie für die Förderung rechenschwacher Schülerinnen und Schüler im Spezifischen verschiedene professionelle Kompetenzen und Wissensbestände (vgl. Kap. 3.2, 3.3 und 3.4). Für erfolgreiches mathematisches Lernen von Schülerinnen und Schülern hat sich als wichtigster Einflussfaktor aufseiten der Lehrperson das mathematikdidaktische Wissen herausgestellt (Hill et al., 2005). Aus der Forschung zu Fort- und Weiterbildungsmassnahmen weiss man zudem, dass professionelle Kompetenzen bzw. Wissensbestände gesteigert werden können (vgl. Kap. 3.4). Dabei hat sich gezeigt, dass dafür fachspezifische Massnahmen, die einen konkreten Bezug zum Curriculum haben und auf evidenzbasierten Erkenntnissen aufbauen, sinnvoll sind (Lipowsky, 2011). Als erfolgreich erwiesen sich zudem Massnahmen, bei denen basierend auf Unterrichtsvideos einerseits das Handeln der Lehrperson und andererseits die Lern- und Verstehensprozesse der Schülerinnen und Schüler reflektiert werden (z.B. Carpenter et al., 1989; Krammer et al., 2012). Zudem sind eine gemeinsame Planung und eine gemeinsame Reflexion von Unterricht mit anderen Lehrpersonen Bestandteil erfolgreicher Coachingansätzen (Staub, 2014).

4.1.3 Offene Fragen

Auf Basis der oben dargestellten, aktuellen Forschungsergebnisse kann die Wirksamkeit von verstehensorientierten Förderansätzen für rechenschwache Schülerinnen und Schüler in der Primarschule, die auf zentrale Inhalte der Mathematik fokussieren und unterrichtliche Massnahmen treffen, die sich als effektiv erwiesen haben (z.B. Einsatz von Veranschaulichungen, Anregen zum Verbalisieren, Geben von gezielten Hinweisen etc.), als bestätigt gelten. Es fällt dabei auf, dass der grösste Teil der einschlägigen Interventionsstudien Kleingruppensettings untersucht hat, wobei die Förderung fast ausschliesslich in separaten Räumen durch speziell geschulte Personen durchgeführt wurde, die nicht an der jeweiligen Schule gearbeitet haben (vgl. Kap. 3.1). Bezüglich sozialer Akzeptanz und mathematischer Entwicklung gibt es jedoch empirische Hinweise sowie theoretische Überlegungen, die für die Wichtigkeit gemeinsamer Lernsituationen sprechen (vgl. Kap. 3.2.2). Zudem kann in diesen Interventionsstudien nicht von einem natürlichen Setting im Sinne eines regulären schulischen Kontextes ausgegangen werden.

Noch weitgehend als ungeklärt gelten muss daher die Frage, inwiefern rechenschwache Schülerinnen und Schüler auch im Rahmen eines regulären Mathematikunterrichts in einer Regelklasse gefördert werden können, und die Förderung somit in einem möglichst natürlichen schulischen Kontext umgesetzt wird. Im Folgenden wird für eine solche mathematische Förderung der Begriff „unterrichtsintegriert“ gewählt. Zudem stellt sich die Frage nach Fortbildungsmassnahmen für Lehrpersonen im Rahmen einer solchen Förderung, in denen die Lehrpersonen zusätzliches Wissen zu Lernprozessen und Lernschwierigkeiten rechenschwacher Schülerinnen und Schüler bzw. zur Förderung dieser Kinder erhalten und in denen die Umsetzung einer solchen Förderung im Unterricht reflektiert wird (vgl. Kap. 3.4). Abschliessend soll noch darauf hingewiesen werden, dass inzwischen verschiedene Faktoren, die die Mathematikleistung beeinflussen, identifiziert worden sind, diesbezüglich aber noch nicht von einem eindeutigen Forschungsstand ausgegangen werden kann (vgl. Kap. 2.3).

In der vorliegenden Arbeit soll ein Beitrag zur Klärung dieser offenen Fragen geleistet werden, indem die Wirksamkeit einer verstehensorientierten Mathematikförderung untersucht wird, die im regulären Mathematikunterricht der Primarschule umgesetzt wird. Im Folgenden soll aufbauend auf den bisher präsentierten Erkenntnissen nochmals zusammenfassend dargestellt werden, wie eine solche Mathematikförderung gestaltet werden sollte.

4.2 (Mathematik)didaktische Folgerungen für eine unterrichtsintegrierte Mathematikförderung in der dritten Klasse der Primarschule

4.2.1 Fokussierung auf zentrale Inhalte

Wie in der vorliegenden Arbeit dargelegt wurde, wird als essenziell erachtet, dass eine mathematische Förderung inhaltspezifisch ist (vgl. Kap. 3.3.1), sich an Modellen der mathematischen Verarbeitung und Entwicklung orientiert (vgl. Kap. 2.2) und empirische Erkenntnisse zu spezifischen Schwierigkeiten rechenschwacher Schülerinnen und Schüler berücksichtigt (vgl. Kap. 2.4). Die mathematischen Inhalte, die aufgrund der bisherigen Überlegungen dieser Arbeit für eine Mathematikförderung als wesentlich erachtet und ins Zentrum der Förderung gestellt werden, werden im Folgenden erläutert. Dazu muss noch angemerkt werden, dass in der dritten Klasse der Zahlenraum auf tausend erweitert wird und zentrale Erkenntnisse und Kompetenzen, die in der ersten und der zweiten Klasse im Hunderterraum erarbeitet wurden (z.B. Stellenwertschreibweise, Bündelungsprinzip, kardinaler und ordinaler Zahlaspekt, Rechenstrategien), erneut thematisiert und auf den höheren Zahlenraum übertragen werden.

Dezimales Stellenwertsystem

Das dezimale Stellenwertsystem wird als eines der zentralsten Konzepte des Arithmetikunterrichts der Primarschule angesehen (z.B. Krauthausen & Scherer, 2007). Untersuchungen weisen zudem darauf hin, dass rechenschwache Schülerinnen und Schüler in diesem Bereich Schwierigkeiten aufweisen (vgl. Kap. 2.4.5). Speziell zu behandelnde Aspekte sind die Stellenwertschreibweise und die Prinzipien des Bündelns und Entbündelns (vgl. Kap. 3.3.2). Diese können in Zusammenhang mit der Erweiterung des Zahlenraums auf tausend nochmals thematisiert und vertieft werden, da u.a. das Prinzip der fortgesetzten Bündelung erst bei grösseren Zahlen in aller Deutlichkeit aufgezeigt werden kann (Scherer & Moser Opitz, 2010). Als geeignete Arbeitsmittel für die Erarbeitung des dezimalen Stellenwertsystems wird das Dienes-

Material in Verbindung mit Stellenwertkarten und Stellenwerttafel empfohlen (vgl. Freesemann, 2014; Van de Walle, 2007). Im Fokus stehen dabei ein strukturorientiertes Verständnis des Stellenwertsystems und der kardinale Zahlaspekt, die in gemeinsamen Lernsituationen mit allen Schülerinnen und Schülern diskutiert und reflektiert werden sollen (vgl. Kap. 3.3.2). Wichtig ist hierzu auch, dass systematische Veränderungen beim Entbündeln (z.B.: Wo verändert sich etwas, wenn ich einen Einer bzw. einen Zehner oder einen Hunderter wegnehmen?) sowie unterschiedliche Bündelungs- und Notationsformen in der Stellenwerttafel thematisiert werden (vgl. Kap. 3.3.2).

Ein weiterer zentraler Aspekt in Zusammenhang mit dem dezimalen Stellenwertsystem ist der ordinale Zahlaspekt und somit ein positionsorientiertes Verständnis von Zahlen (vgl. Freesemann, 2014). Als Arbeitsmittel eignen sich in diesem Zusammenhang die Hunderterkette und der Zahlenstrahl (vgl. Kap. 3.3.2). In Anlehnung an Höhtker und Selter (1995) empfiehlt sich eine schrittweise Erarbeitung ausgehend von der Hunderterkette hin zum skalierten Zahlenstrahl, wobei Orientierungsübungen wie „Zahlen finden", „Orte finden" und „Nachbarzahlen finden" einen wichtigen Beitrag dazu leisten können, damit Schülerinnen und Schüler einen inneren, mentalen Zahlenstrahl aufbauen und sich im Zahlenraum besser orientieren können. Gemeinsam zu diskutieren sind zudem geeignete Orientierungsstützpunkte am Zahlenstrahl (vgl. Kap. 3.3.2).

Halbschriftliche Rechenstrategien für die Addition und die Subtraktion

Wie verschiedene Untersuchungen zeigen, weisen rechenschwache Schülerinnen und Schüler einen eingeschränkten Strategiengebrauch beim Kopfrechnen auf (vgl. Kap. 2.4.3). Mit der Erweiterung des Zahlenraums auf tausend wird daher als wichtig angesehen, dass diese Schülerinnen und Schüler eine für sie geeignete halbschriftliche Rechenstrategie für die Addition und die Subtraktion erarbeiten können (vgl. Kap. 3.3.2). Halbschriftliche Rechenstrategien sind im Vergleich zum Kopfrechnen mit einer geringeren Belastung des Arbeitsgedächtnisses verbunden, weswegen sie bei Rechenschwäche besonders geeignet sind (z.B. Scherer & Moser Opitz, 2010; vgl. 2.3.1).

Für rechenschwache Schülerinnen und Schüler empfehlen sich insbesondere die Rechenstrategien *Schrittweise* und *Stellenwerte extra*, wobei bei Letzterer einige Tücken bei der Subtraktion beachtet werden müssen (vgl. Kap. 3.3.2). Für die Erarbeitung der Rechenstrategien eignet sich das Dienes-Material in Verbindung mit den Stellenwertkarten sowie der Rechenstrich zum Protokollieren der einzelnen Rechenschritte (vgl. Kap. 3.3.2). Als wesentlich wird dabei erachtet, dass eigene Rechen- bzw. Denkwege verbalisiert werden, verschiedene Strategien und Notationsformen miteinander verglichen werden und eine gemeinsame Reflexion dieser Aspekte im Klassenverband erfolgt (z.B. Scherer & Moser Opitz, 2010). Für rechenschwache Schülerinnen und Schüler sollte dabei das Finden und Erarbeiten der für sie passendsten Rechenstrategie im Fokus stehen (Schmassmann & Moser Opitz, 2008b).

Mathematisieren

Für die mathematische Entwicklung als wichtig erachtet werden Mathematisierungskompetenzen. D.h., Schülerinnen und Schüler müssen Einsicht in die mathematischen Handlungen haben, die hinter den vier Grundoperationen stehen, und die Verbindung zur Gleichungsschreibweise herstellen können (vgl. Kap. 3.3.2). In diesem Bereich weisen rechenschwache Schülerinnen und Schüler Schwierigkeiten auf (vgl. Kap. 2.4.4). Zentral für eine Förderung ist das Veranschaulichen der vier Grundoperationen mit Material, Rechengeschichten und Bildern (z.B. Häsel-Weide, 2016; vgl. Kap. 3.3.2). Dabei ist wichtig, mathematische Handlungen in Worte zu fassen, mit der formalen Schreibweise zu verbinden und zwischen den verschiedenen Repräsentationsformen hin und her zu wechseln und diese miteinander zu verknüpfen (z.B. Krauthausen & Scherer, 2007). Zudem muss thematisiert werden, dass Darstellungen auch mehrdeutig sein können (z.B. Scherer & Moser Opitz, 2010), und die Lehrpersonen müssen beachten, dass bei der Multiplikation und der Division verschiedene Modellvorstellungen berücksichtigt werden müssen (vgl. Kap. 3.3.2). Empfohlen werden auch Übungen zum Aufbau von Grundvorstellungen, bei denen ausgehend von konkreten mathematischen Handlungen, die am Material ausgeführt werden, schrittweise eine Bearbeitung auf symbolischer Ebene erarbeitet wird (Wartha & Schulz, 2011; vgl. Kap. 3.3.2).

Flexible Zählkompetenz

Zählen gehört zu den mathematischen Vorläuferfertigkeiten und ist daher eine Kompetenz, die für die mathematische Entwicklung von grosser Bedeutung ist (z.B. Krajewski & Schneider, 2006; vgl. Kap. 2.2). Zudem zeigen empirische Untersuchungen, dass auch in höheren Klassenstufen rechenschwache Schülerinnen und Schüler noch Schwierigkeiten aufweisen, wenn es um komplexere Zählkompetenzen geht (z.B. Moser Opitz, 2013). Für eine Mathematikförderung in der dritten Klasse empfehlen sich daher Übungen zum verbalen Zählen, wobei das Zählen vorwärts und rückwärts in Schritten grösser als eins von verschiedenen Startzahlen aus geübt werden sollte (vgl. Kap. 3.3.2).

Aufbau von tragfähigem Faktenwissen

Wie in dieser Arbeit aufgezeigt wurde, haben rechenschwache Schülerinnen und Schüler grosse Mühe beim Faktenabruf (vgl. Kap. 2.4.3). Es wird daher als wichtig erachtet, rechenschwache

Schülerinnen und Schüler beim Aufbau von tragfähigem Faktenwissen aktiv zu unterstützen. Dabei soll der Fokus jedoch nicht auf ein Auswendiglernen gelegt, sondern eine verstehensorientierte Herangehensweise gewählt werden (z.B. Krauthausen & Scherer, 2007). Hierfür eignen sich insbesondere operativ strukturierte Aufgaben im Sinne von Wittmann (1992), bei denen systematische Variationen zwischen den Aufgaben helfen, Strukturzusammenhänge und Gesetzmässigkeiten zu erkennen und zu verinnerlichen. Wichtig für den Faktenabruf ist auch der Aufbau von Einsicht in Teile-Ganzes-Beziehungen, was mit Aufgaben zum Halbieren, Verdoppeln und Ergänzen erfolgen kann (z.B. Häsel-Weide, 2016).

Ablösung vom zählenden Rechnen

Da auch in der dritten Klasse von rechenschwachen Schülerinnen und Schülern noch gehäuft zählende Rechenstrategien eingesetzt werden (vgl. Kap. 2.4.3), sollten bei einer Mathematikförderung an verschiedenen Stellen Massnahmen eingebaut werden, die bei der Ablösung von zählenden Rechenstrategien helfen (vgl. Häsel-Weide, 2016). Neben der Förderung flexibler Zählkompetenzen, der Förderung von Einsicht in Teile-Ganzes-Beziehungen und der Förderung von Grundvorstellungen zu den vier Grundoperationen werden daher zusätzlich auch Massnahmen für die Förderung einer strukturierten Anzahlerfassung eingebaut (z.B. beim Verwenden des Dienes-Materials auffordern, dass dieses so gelegt wird, dass es auf einen Blick erfasst werden kann; vgl. Kap. 3.3.2).

4.2.2 Einsatz von geeigneten Veranschaulichungen und Hilfsmitteln

Diese Arbeit hat verdeutlicht, dass für einen verstehensorientierten Mathematikunterricht unter Berücksichtigung rechenschwacher Schülerinnen und Schüler der Einsatz geeigneter Veranschaulichungen und Hilfsmittel als wichtig erachtet wird (vgl. z.B. Kap. 3.3.1). Damit kann der Aufbau mentaler Vorstellungsbilder gefördert und gleichzeitig auch das Arbeitsgedächtnis entlastet werden (vgl. z.B. Scherer & Moser Opitz, 2010; W. Schneider et al., 2013). Die Auswahl der Arbeitsmittel bzw. Veranschaulichungen muss dabei mathematikdidaktisch begründet sein (vgl. Krauthausen & Scherer, 2007). Zudem wird für rechenschwache Schülerinnen und Schüler als unabdingbar angesehen, dass eine explizite Einführung der Materialien erfolgt, in der die Struktur und der Verwendungszweck der Veranschaulichungen und Hilfsmittel mit den Schülerinnen und Schülern explizit erarbeitet werden (vgl. die Ausführungen in Kap. 3.3.1). Für die unterrichtsintegrierte Mathematikförderung der vorliegenden Untersuchung wurden folgende Materialien gewählt, die jeweils mit der ganzen Klasse erarbeitet werden:

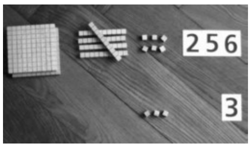
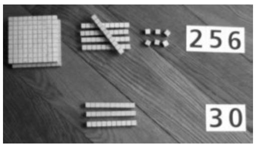
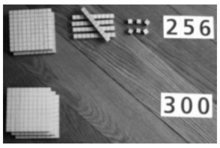
- Hunderterkette und Zahlenstrahl: Darstellung des ordinalen Zahlaspekts bzw. des positionsorientierten Zahlenverständnisses, unterstützt den Aufbau eines mentalen Zahlenstrahls und die Orientierung im Zahlenraum,
- Dienes-Material: Darstellung des kardinalen Zahlaspekts, eignet sich für die Erarbeitung des dezimalen Stellenwertsystems (Prinzipien des Bündelns, des Entbündelns und der Stellenwertschreibweise). Wichtig ist, dass eine Verbindung mit der formalen Schreibweise erfolgt (→ Stellenwertkarten, Stellenwerttafel),
- Stellenwertkarten und Stellenwerttafel: Veranschaulichung des Prinzips der Stellenwertschreibweise (Bedeutung der Position einer Ziffer in einer Zahl), kann in Verbindung mit dem Dienes-Material eingesetzt werden, um eine Verknüpfung mit formaler Schreibweise herzustellen,
- Rechenstrich: Kann zum Protokollieren von Rechenaufgaben eingesetzt werden (Addition und Subtraktion), es können unterschiedlich anspruchsvolle Rechenschritte durchgeführt werden (z.B. zuerst Einer, dann Zehner, dann Hunderter), und das Anwenden komplexerer Rechenschritte kann erarbeitet werden (z.B. Einer und Zehner zusammenfassen und gemeinsam addieren).

Damit der handelnde Umgang mit Materialien auch tatsächlich zum Aufbau von mentalen Vorstellungsbildern führt, ist es wichtig, dass die Schülerinnen und Schüler immer wieder zu einer Versprachlichung der mathematischen Handlungen am Material aufgefordert werden, da so eine Verknüpfung zwischen Handlungen und formalen Schreibweisen eingeleitet werden kann (vgl. Kap. 3.3.1). Längerfristig soll zudem eine Ablösung vom Material angeregt werden, wie sie beispielsweise von Wartha und Schulz (2011) vorgeschlagen wird.

4.2.3 Produktives Üben

Ein weiterer zentraler Aspekt für eine verstehensorientierte Mathematikförderung ist eine produktive Übungspraxis (vgl. Kap. 3.1.1 und 3.3.1). Dabei wird als wichtig erachtet, dass das Üben so gestaltet wird, dass Einsicht in Strukturzusammenhänge und Gesetzmässigkeiten ermöglicht wird, anstatt dass ein zu früher Fokus auf eine Automatisierung gelegt wird (z.B. Krauthausen & Scherer, 2007). Ein Übungsformat, das für den Aufbau von Verständnis und Einsicht als besonders geeignet angesehen und daher im Rahmen der unterrichtsintegrierten Förderung an verschiedenen Stellen eingesetzt wird, sind operativ strukturierte Aufgaben nach Wittmann (1992). Mittels einer systematischen Variation der Aufgabenbeispiele werden die Schülerinnen und Schüler durch solche Aufgaben in Phasen des individuellen Übens darin unterstützt, Strukturzusammenhänge zu erkennen und ein tieferes Verständnis für die zu lernenden Konzepte bzw. Kompetenzen aufzubauen, da die Aufgabenstruktur dabei hilft den Kern der Sache zu erkennen. Solche Übungsformate können in verschiedenen Phasen des Lernprozesses eingesetzt werden (vgl. Wittmann, 1992). Materialgestützt eignen sie sich insbesondere auch unter Berücksichtigung rechenschwacher Lernenden für die Erarbeitung von prozeduralen Kompetenzen (vgl. Tabelle 3).

Tabelle 3: Beispiele für operativ strukturierte Aufgaben für unterschiedliche Phasen des Lernprozesses

Materialgestützte Aufgabenbeispiele	256 + 3 = ___ 	256 + 30 = ___ 	256 + 300 = ___ 
Aufgabenbeispiele zum Automatisieren	80 + 7 = ___ 70 + 17 = ___ 70 + 27 = ___	30 + 21 = ___ 30 + 31 = ___ 40 + 41 = ___	50 + 29 = ___ 40 + 29 = ___ 30 + 19 = ___

Ebenso kann aber auch eine Automatisierung von Zahlenfakten mit produktiven Übungsformaten gefördert werden (vgl. Tabelle 3). Wichtig zu beachten ist in diesem Zusammenhang aber, dass eine Automatisierung erst erfolgen sollte, wenn Einsicht und Verständnis aufgebaut sind (vgl. Kap. 3.3.1). Wenn Schülerinnen und Schüler diesbezüglich noch Schwierigkeiten haben, sollten Aufgaben zunächst unter Zuhilfenahme von Arbeitsmitteln bearbeitet werden.

4.2.4 Gemeinsame Lernsituationen

Das Ermöglichen von gemeinsamen Lernsituationen wird aus entwicklungstheoretischen, gemeinsamkeitsstiftenden und empirischen Gründen als wichtig erachtet (vgl. Kap. 3.2.2). Um der Gefahr der Vereinzelung entgegenzuwirken, sollte daher eine unterrichtsintegrierte Förderung geplant werden, in der Schülerinnen und Schüler mit unterschiedlichen Lernvoraussetzungen an den gleichen Themen arbeiten. Dafür müssen Inhalte so aufbereitet werden, dass bedeutende Konzepte und Kompetenzen der dritten Klasse verstehensorientiert erarbeitet werden können und schwache Rechnerinnen und Rechner gleichzeitig die Gelegenheit erhalten, Lücken im Lernstoff zu schliessen (vgl. bisherige Ausführungen in Kap. 4.2). Dabei ist zu beachten, dass für eine unterrichtsintegrierte Mathematikförderung zusätzlich zu gemeinsamen Lernsituationen auch Unterrichtsphasen notwendig sind, in denen Schülerinnen und Schüler stärker individualisiert arbeiten können (vgl. Scherer & Moser Opitz, 2010).

Im Rahmen einer unterrichtsintegrierten Mathematikförderung können verschiedene Unterrichtsphasen genutzt werden, um sowohl gemeinsame Lernsituationen als auch Gelegenheiten für individuelles Arbeiten zu ermöglichen (z.B. Wocken, 1998; vgl. Kap. 3.2.2). Für gemeinsame Lernsituationen eignen sich beispielsweise Einführungs- und Reflexionsphasen. In Einführungsphasen können gemeinsame Lernsituationen in Form eines Klassendiskurses umgesetzt werden, um neue Konzepte bzw. Kompetenzen gemeinsam zu erarbeiten und unterschiedliche Perspektiven bzw. Herangehensweisen der Schülerinnen und Schüler für den Verständnisaufbau zu nutzen (vgl. Kap. 3.2.1). Auch Reflexionsphasen eignen sich sehr gut für gemeinsame Lernsituationen, da so individuelle Erkenntnisprozesse im gemeinsamen Gespräch aufgegriffen, diskutiert und reflektiert werden können. Damit alle Schülerinnen und Schüler trotz unterschiedlichen Lernvoraussetzungen in Einführungs- und Reflexionsphasen geeignet einbezogen werden können, empfiehlt sich für diese Sequenzen eine Differenzierung durch Aufgabenbeispiele (z.B. unterschiedliche Zahlenräume, Verwendung von Materialien; vgl. Scherer & Moser Opitz, 2010). Für stärker individualisierte Sequenzen im Rahmen einer unterrichtsintegrierten Förderung können hingegen individuelle Arbeitsphasen genutzt werden. In diesen können gemeinsam erarbeitete Konzepte und Kompetenzen unter Verwendung produktiver Übungsformate individuell vertieft und die Schülerinnen und Schüler durch eine adaptive Lernbegleitung individuell unterstützt werden (vgl. Kap. 3.2 und 3.3). Dabei sollte die Lehrperson an konkreten Schwierigkeiten bzw. Denk- und Verstehensprozessen der einzelnen Kinder ansetzen, um eine optimale Lernunterstützung zu ermöglichen (vgl. Kap. 3.2.1). Unterricht soll im Rahmen einer unterrichtsintegrierten Mathematikförderung jeweils ausgehend von den spezifischen Schwierigkeiten und Anforderungen rechenschwacher Schülerinnen und Schüler unter Berücksichtigung didaktischer und mathematikdidaktischer Überlegungen geplant werden (vgl. Scherer & Moser Opitz, 2010).

4.2.5 Adaptive Lernbegleitung

Wie empirische Untersuchungen zeigen, ist ein hoher Grad an Strukturierung, wie sie bei Ansätzen der direkten Instruktion zum Einsatz kommt, ein wichtiger Aspekt erfolgreicher Förderansätze für rechenschwache Schülerinnen und Schüler (vgl. Kap. 3.1.1). Allerdings wird an entsprechenden Instruktionsansätzen eine zu geringe Verstehensorientierung bemängelt und deren Einsatz in heterogenen Lerngruppen als problematisch angesehen (z.B. Montague, 2011). Alternativ empfiehlt sich für eine unterrichtsintegrierte Mathematikförderung daher der Ansatz der adaptiven Lernbegleitung (z.B. Beck et al., 2008; Parsons et al., 2018; vgl. Kap. 3.2.1). Eine adaptive Lernbegleitung zielt darauf ab, eine bestmögliche Passung zwischen den unterrichtlichen Handlungen der Lehrperson und den Voraussetzungen der Schülerinnen und Schüler herzustellen, sodass individuelle Lernprozesse optimal unterstützt werden (z.B. Leiss, 2010). Der Ansatz der adaptiven Lernbegleitung bietet zudem Hinweise, Strategien sowie konkrete

Massnahmen, die sich für eine Umsetzung im Rahmen einer mathematischen Förderung eignen und eine strukturierte Lernbegleitung ermöglichen (vgl. Pfister et al., 2015; van de Pol et al., 2010). Lehrpersonen sollen demnach auf inhaltlicher und prozeduraler Ebene gezielte Hinweise, konkrete Anweisungen und ergänzende Erklärungen geben, die die Schülerinnen und Schüler zum Kern der Sache führen. Auch bei der Auswahl der Aufgaben sollte darauf geachtet werden, dass die Lernenden darin unterstützt werden sich zentrale Erkenntnisse zu erarbeiten (siehe hierzu auch Kap. 4.2.3). Zudem sollen Lösungsstrategien im Sinne eines Lernens am Modell vorgezeigt und Feedbacks zu individuellen Lernprozessen gegeben werden. Ein weiterer wichtiger Aspekt ist zudem die Diskursanregung, die durch das Stellen von Fragen auf eine Versprachlichung und eine gemeinsame Reflexion abzielt. Zudem sollen herausfordernde Aufgaben und Fragen zu einem hohen Ausmass an kognitiver Aktivierung führen. Individuelle Unterstützung muss ausserdem an individuellen Schwierigkeiten und Fehlern ansetzen und unter Zuhilfenahme geeigneter Arbeitsmittel ein verstehensorientiertes Lernen ermöglichen (vgl. Kap. 3.2.1).

Je nach Unterrichtsphase können verschiedene Aspekte einer adaptiven Lernbegleitung zum Einsatz kommen:

- In Einführungsphasen soll die Lehrperson an ausgewählten Aufgaben gezielt Hinweise geben und Fragen stellen, die zum Kern der Sache führen. Prozeduren können gemeinsam erarbeitet und beispielhaft vorgezeigt werden, wobei konkrete Anweisungen zu den einzelnen Schritten gegeben und bei Bedarf auch ergänzende Erklärungen angefügt werden sollten,
- In Arbeitsphasen muss die Lehrperson individuelle Unterstützung bieten, indem sie gezielt Fragen stellt, Fehler als Ausgangspunkt für weitere Hinweise und Erklärungen nutzt und Feedback zu den individuellen Lernprozessen gibt,
- In Reflexionsphasen steht die Diskursanregung im Mittelpunkt. Die Schülerinnen und Schüler sollen sich über das Gelernte austauschen und in eine gemeinsame Reflexion treten, wobei die Lehrperson auch hier gezielt strukturieren muss.

4.2.6 Einordnung der unterrichtsintegrierten Mathematikförderung im Angebots-Nutzungs-Modell

In der aktuellen Unterrichtsforschung wird Unterricht meistens im Sinne des Angebots-Nutzungs-Modells betrachtet (vgl. Seidel, 2014). Dieses Modell basiert auf Arbeiten von Fend (1981) sowie von Helmke und Weinert (1997) und wurde in den vergangenen Jahren vielfach weiterentwickelt und ergänzt (Seidel, 2014). Unterricht wird im Sinne dieses Modells (vgl. Abbildung 7) als Angebot verstanden, das in Abhängigkeit von Merkmalen der Lehrperson (professionelle Kompetenzen, allgemeine Charakteristika) und unter Einfluss verschiedener Kontextbedingungen (Bildungssystem, Klasse, Schule, Kollegium und Fach) unterschiedliche Qualität annehmen kann und aufseiten der Schülerinnen und Schüler zu unterschiedlichen individuellen Lernaktivitäten führt. Dabei ist die Möglichkeiten der Nutzung des Unterrichtsangebotes von individuellen Voraussetzungen (z.B. Intelligenz, Vorwissen, Motivation) abhängig, die wiederum von der individuellen Lernumwelt (Familie, Gleichaltrige, Medien) beeinflusst werden (Seidel, 2014). Der Output dieses Angebots-Nutzungs-Modell sind die Lernergebnisse, die sich aus fachlichen (Wissen, Können, Überzeugungen, Interesse) und überfachlichen Kompetenzen (z.B. Lernstrategien) zusammensetzen (Seidel, 2014).

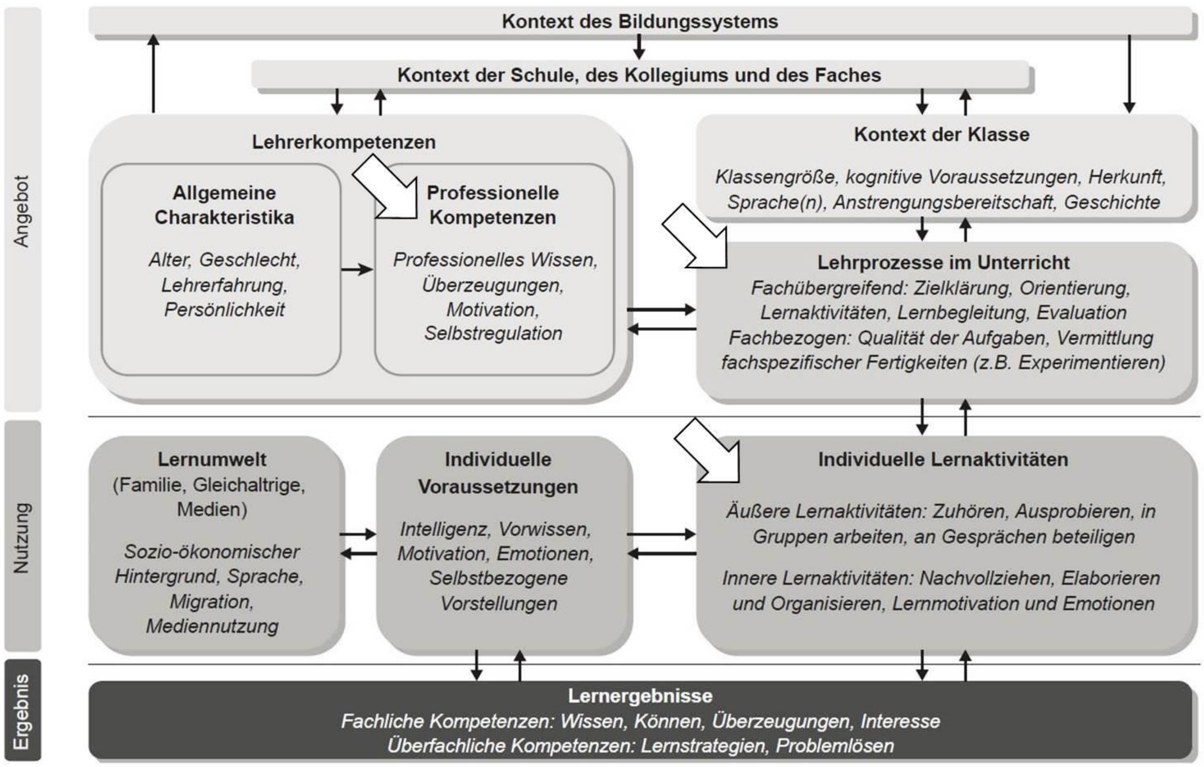


Abbildung 7: Einordnung der unterrichtsintegrierten Mathematikförderung (weisse Pfeile) im Angebots-Nutzungs-Modell von Fend (1981) und Helmke und Weinert (1997) gemäss der Modelldarstellung von Seidel (2014, S. 858)

Eine unterrichtsintegrierte Mathematikförderung, wie sie in der vorliegenden Arbeit untersucht wird, setzt an drei verschiedenen Punkten dieses Angebots-Nutzungs-Modells an (vgl. weisse Pfeile in Abbildung 7). Erstens sollen auf Angebotsseite die *professionellen Kompetenzen* der Lehrpersonen durch schriftliche Anleitungen und Materialien, die ihnen für die unterrichtsintegrierte Mathematikförderung zur Verfügung gestellt werden, sowie durch zusätzliches Wissen zu rechenschwachen Kindern und deren Förderung, das im Rahmen von Professionalisierungsmassnahmen vermittelt wird, gesteigert werden (vgl. Kap. 3.4). Zweitens soll auf Angebotsseite die Qualität der *Lehrprozesse* durch die zur Verfügung gestellten schriftlichen Anleitungen und Materialien und die darin enthaltenen Hinweise zu einer adaptiven Lernbegleitung gestärkt werden (vgl. Kap. 3.2 und 3.3). Drittens soll als Folge davon auf Nutzungsseite die Qualität der *individuellen Lernaktivitäten* gesteigert werden (vgl. Kap. 3.2 und 3.3). Schliesslich sollen diese drei Aspekte dazu führen, dass die *Lernergebnisse* der Schülerinnen und Schüler im Fach Mathematik ebenfalls verbessert werden (vgl. Kap. 3.2 und 3.3).

5. Eigene Untersuchung

Nachdem im letzten Kapitel Schlussfolgerungen für eine unterrichtsintegrierte Mathematikförderung gezogen wurden und die Förderung im Angebots-Nutzungs-Modell (Fend, 1981; Helmke & Weinert, 1997; Seidel, 2014) eingeordnet worden ist, wird im Folgenden die Untersuchung, auf der diese Arbeit beruht, detailliert vorgestellt. Ausgehend von den Fragestellungen dieser Arbeit, wird das Design der Interventionsstudie besprochen und die Intervention erläutert. Zudem werden die eingesetzten Messinstrumente thematisiert, und eine ausführliche Beschreibung der Gesamtstichprobe wird präsentiert. Abschliessend wird auf die Auswertungsmethoden eingegangen, mit denen die Fragestellungen der Untersuchung beantwortet werden.

5.1 Fragestellungen und Hypothesen

Basierend auf den bisherigen Erkenntnissen dieser Arbeit sowie den offenen Fragen, die im Kapitel 4 zusammengefasst wurden, werden in der vorliegenden Untersuchung die folgenden Fragestellungen analysiert und beantwortet:

1. Fragestellung:
Welche erhobenen Merkmale auf Individual- und auf Klassenebene haben einen Einfluss auf die Mathematikleistung?

2. Fragestellung:
Kann durch eine von der Lehrperson durchgeführte, unterrichtsintegrierte Förderung im Fach Mathematik, die auf zentrale arithmetische Inhalte fokussiert und Aspekte wirksamen Mathematikunterrichts berücksichtigt, im Vergleich zu einem regulären Mathematikunterricht eine Verbesserung der Mathematikleistungen erreicht werden?

Hypothese:
Schülerinnen und Schüler, deren Lehrpersonen eine spezifische, unterrichtsintegrierte Förderung im Fach Mathematik durchgeführt haben, die auf zentrale arithmetische Inhalte fokussiert und Aspekte eines wirksamen Mathematikunterrichts berücksichtigt, machen grössere Leistungsfortschritte als Schülerinnen und Schüler, deren Lehrpersonen keine spezifische Förderung eingesetzt haben.

3. Fragestellung:
Hat die Form, mit der die Lehrpersonen angeleitet werden (schriftliche Materialien und Anleitungen vs. intensivere Begleitung durch Fortbildungstreffen), einen Einfluss auf die Leistungsfortschritte der Schülerinnen und Schüler?

Hypothese:
Schülerinnen und Schüler, deren Lehrpersonen eine intensivere Begleitung in Form von Fortbildungstreffen erhalten haben, können stärker von der Förderung profitieren als Schülerinnen und Schüler, deren Lehrpersonen nur die schriftlichen Materialien und Anleitungen für die unterrichtsintegrierte Förderung erhalten haben.

5.2 Untersuchungsdesign

Um die Frage nach der Wirksamkeit einer unterrichtsintegrierten Mathematikförderung, die auf zentrale Inhalte der Arithmetik fokussiert, beantworten zu können, wurde ein quasi-experimentelles Untersuchungsdesign mit zwei Interventionsgruppen (Gruppe$^{\text{Begl}}$ und Gruppe$^{\text{Mat}}$) und einer Kontrollgruppe (Gruppe$^{\text{Kontr}}$) sowie vier Messzeitpunkten, an denen die Mathematikleistung erfasst wurde, gewählt (vgl. Abbildung 8). Die Klassen wurden zufällig auf die drei Stichprobengruppen aufgeteilt. Es wurde jedoch darauf geachtet, dass in allen drei Stichprobengruppen Klassen aus unterschiedlich stark urbanisierten Regionen der Schweiz vertreten sind.

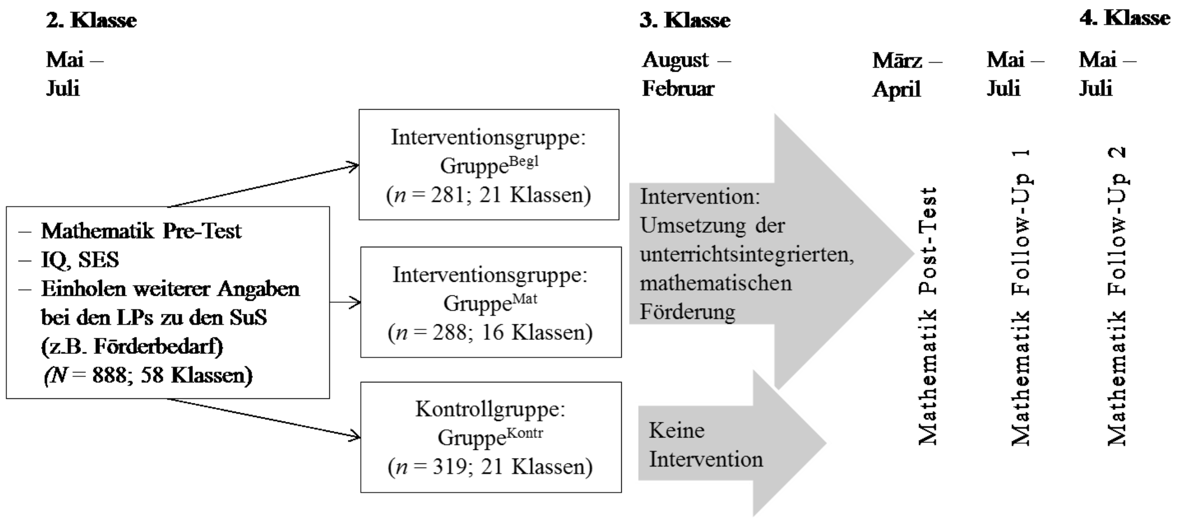


Abbildung 8: Forschungsdesign (IQ = Intelligenz, SES = sozioökonomischer Status, LP = Lehrperson, SuS = Schülerinnen und Schüler)

Ein quasi-experimentelles Untersuchungsdesign ermöglicht Aussagen zur Wirksamkeit einer Intervention, da neben den Interventionsgruppen auch eine Kontrollgruppe untersucht wird und so spezifische Unterrichtsmassnahmen mit einem Unterricht ohne besondere Fördermassnahmen verglichen werden können (Gniewosz, 2015). Zusätzlich wurde mit zwei unterschiedlichen Interventionsformen untersucht, ob eine intensivere Begleitung der Lehrpersonen (Gruppe$^{\text{Begl}}$) durch das Projektteam im Vergleich zu schriftlichen Materialien und Anleitungen

(Gruppe$^{\text{Mat}}$) für die Lehrpersonen zusätzlich einen positiven Effekt auf die Mathematikleistungen der Schülerinnen und Schüler hat.

Es handelt sich in der vorliegenden Untersuchung um eine Selbstselektionsstichprobe (Döring & Bortz, 2016), bei der sich Lehrpersonen aus verschiedensten Gründen für eine Teilnahme mit ihrer Klasse entschieden haben. Die Zuteilung der Klassen auf die Stichprobengruppen erfolgte jedoch randomisiert. Für die Gewinnung der Gesamtstichprobe wurde eine Website mit Informationen zum Projekt und zu den Anmeldeformalitäten eingerichtet sowie Flyer erstellt. Diese wurden an dafür Verantwortliche aus der Bildungsverwaltung, Schulleitungen und Lehrpersonen aus der Deutschschweiz verteilt, sodass sich interessierte Personen der gesuchten Klassenstufe melden konnten. Bei Untersuchungen im realen Schulkontext ist es meistens nicht möglich, eine Zufallsstichprobe zu ziehen, da man immer auf eine freiwillige Teilnahme auf Ebene der Kantone, Gemeinden, Schulhäusern, Lehrpersonen und der Kinder bzw. deren Erziehungsberechtigten angewiesen ist.

Vor Untersuchungsbeginn wurde den Lehrpersonen mitgeteilt, welcher Stichprobengruppe sie zugeteilt wurden. Die Lehrpersonen der beiden Interventionsgruppen (Gruppe$^{\text{Begl}}$ und Gruppe $^{\text{Mat}}$) haben die ausführlichen Unterrichtsmaterialien und -unterlagen (vgl. Kap. 5.3.1) an einer Einführungsveranstaltung erhalten, die nach den Sommerferien am Anfang der dritten Klasse an verschiedenen Orten in der Deutschschweiz durchgeführt wurde. Die Lehrpersonen der Gruppe$^{\text{Begl}}$ haben während der Interventionszeit zusätzlich an zwei Fortbildungstreffen teilgenommen und dadurch eine intensivere Begleitung erhalten (vgl. Kap. 5.3.2). Das erste Treffen fand zwischen Oktober und November an verschiedenen Orten in der Deutschschweiz statt, das zweite im Januar.

Das Förderprogramm wurde direkt im Anschluss an die Einführungsveranstaltung im ersten Halbjahr der dritten Klasse in allen Interventionsklassen von den Regellehrpersonen und teilweise mit Unterstützung der schulischen Heilpädagoginnen und Heilpädagogen eingesetzt. Es erstreckte sich über eine Zeitdauer von 20 Wochen. Um die Umsetzung der Förderung in den einzelnen Interventionsklassen überprüfen zu können, füllte jede Regellehrperson in regelmässigen Abständen ein Formular aus und gab an, welche Elemente der Förderung umgesetzt werden konnten und welche Schwierigkeiten es bei der Umsetzung gab.

Eine erste Erfassung der Mathematikleistung, des IQs und des SES sowie die Erhebung zusätzlicher Angaben zu den Schülerinnen und Schülern bei den Lehrpersonen (z.B. Angaben zu Förderbedarf, Einschränkungen im Deutschverstehen, Geburtsdatum etc.; vgl. Kap. 5.4.4), wurden vor der Intervention am Ende der zweiten Klasse im Rahmen des Pre-Tests durchgeführt (vgl. Abbildung 8). Direkt im Anschluss an die Intervention wurde die Mathematikleistung im dritten Quartal der dritten Klasse erneut erfasst (Post-Test). Um auch längerfristige Wirkungen untersuchen zu können, wurde zudem am Ende der dritten Klasse (Follow-up 1) sowie knapp eineinhalb Jahre nach der Intervention am Ende der vierten Klasse nochmals eine Testung der Mathematikleistungen durchgeführt (Follow-up 2). Wenn Kinder an einzelnen Erhebungsterminen krank waren, wurde nach Möglichkeit eine Nachtestung durchgeführt.

Zusätzlich wurde beim Post-Test in einer explorativ angelegten, schriftlichen Befragung auch das Wissen der Regellehrpersonen zu Rechenschwäche erhoben (Symptome, Ursachen, geeignete Fördermöglichkeiten und Förderprogramme) und im Rahmen einer unveröffentlichten Masterarbeit durch die Verfasserin der vorliegenden Arbeit ausgewertet. Um eine möglichst hohe Beteiligung an der Befragung zu ermöglichen, wurden diese Daten jedoch anonym erfasst, sodass diese für die Analysen dieser Arbeit nicht berücksichtigt werden können. Zudem wurde in allen Interventionsklassen von der Projektmitarbeiterin Mirjam Pfister je eine Schulstunde während der Interventionsphase videografiert und die Daten mithilfe eines hochinferenten Ratings bezüglich der Adaptivität der Lernbegleitung analysiert (vgl. Pfister et al., 2015; Pfister, 2016). Die Videos stellen eine weitere wichtige Quelle dar, mit der die Umsetzung der Förderung in den Interventionsklassen kontrolliert werden konnte. Da jedoch für die Kontrollklassen keine Videos vorliegen, können die Daten aus der Videoanalyse nicht für die Beantwortung der Forschungsfragen genutzt werden.

5.3 Intervention

5.3.1 Beschreibung der schriftlichen Anleitungen und Materialien

Im Folgenden werden die schriftlichen Anleitungen und Materialien, die die Lehrpersonen für die unterrichtsintegrierte Förderung erhalten haben, vorgestellt und im Anschluss daran aufgezeigt, wie die intensivere Begleitung der Lehrpersonen der Gruppe^{Begl} gestaltet wurde. Die Ausführungen in diesem Kapitel basieren auf dem aktuellen Stand empirischer Erkenntnisse und didaktischer sowie mathematikdidaktischer Überlegungen, wie sie in dieser Arbeit bisher aufgezeigt wurden. Für die konkrete didaktische und mathematikdidaktische Ausgestaltung der einzelnen Unterrichtseinheiten, die im Folgenden vor allem inhaltlich erläutert werden, wird auf die Ausführungen im Kapitel 4.2 verwiesen.

Aufbau des Förderkonzepts

Die Lehrpersonen haben vom Projektteam schriftliche Anleitungen und Materialien für die unterrichtsintegrierte Förderung zur Verfügung gestellt bekommen.[12] In allen Schulklassen, die an der Untersuchung teilgenommen haben, wurde im regulären Mathematikunterricht entweder mit dem Schweizer Zahlenbuch (3. Klasse: Hengartner & Wieland, 2008) oder dem Zürcher Mathematiklehrmittel (3. Klasse: Keller & Noelle Müller, 2012) gearbeitet. Entsprechend wurden zwei unterschiedliche Versionen der schriftlichen Anleitungen erstellt, die auf das jeweilige Mathematiklehrmittel der dritten Klasse abgestimmt waren. Eine Auflistung der Unterlagen können der Tabelle 4 entnommen werden.

Tabelle 4: Schriftliche Anleitungen und Materialien für die unterrichtsintegrierte Mathematikförderung

Unterrichtseinheiten:
- Dezimalsystem
- Zahlenstrahl
- Addition und Subtraktion

Karteien:
- Zählkartei
- Kopfrechenkartei
- Unterrichtsangebot zum Mathematisieren

Schriftliche Hinweise zur Durchführung:
- Aufbau und zeitlicher Ablauf der Förderung
- Einsatz der Karteien
- Massnahmen und Strategien für eine adaptive Lernbegleitung

Materialien:
- Dienes-Material
- Hunderterkette
- Stellenwertkarten

Unterrichtseinheiten

Der Hauptteil der Förderung bestand aus drei Unterrichtseinheiten zu den Themen Dezimalsystem, Zahlenstrahl sowie Addition und Subtraktion. Diese Unterrichtseinheiten umfassten jeweils mehrere Lektionen und enthielten ausführliche Anweisungen für die Gestaltung der einzelnen Lektionen. Jede Unterrichtseinheit begann mit einer Standortbestimmung, sodass die Förderung basierend auf dem jeweiligen Vorwissen individualisiert werden konnte.

[12] Die schriftlichen Anleitungen und Materialien wurden von der Projektmitarbeiterin Lis Reusser in Zusammenarbeit mit Prof. Dr. Elisabeth Moser Opitz entwickelt.

In der *Unterrichtseinheit „Dezimalsystem“* wurde eine Vorstellung für die Zahlen bis 1000 aufgebaut und die Struktur des dezimalen Stellenwertsystems erarbeitet. Der kardinale Zahlaspekt stand dabei im Fokus, und es wurde besonderes Gewicht auf das Bündelungsprinzip und die Stellenwertschreibweise gelegt. Die detaillierten Ziele der einzelnen Bausteine der Unterrichtseinheit „Dezimalsystem“ können der Tabelle 5 entnommen werden.

Tabelle 5: Übersicht über die Ziele der einzelnen Bausteine aus der Unterrichtseinheit „Dezimalsystem“

Baustein	**Ziele**
Zählen und Bündeln	Bündeln als effektive Strategie kennen
	Erkennen, dass unser Zahlensystem auf Zehnerbündelung beruht
	Wiederholung der Zahlenschreib- und der Sprechweise zweistelliger Zahlen
Würfelmaterial	Kennen des Dienes-Materials, Einheiten benennen können (Einer, Zehner, Hunderter, Tausender)
	Erkennen und wissen, dass zehn Einerwürfel einem Zehnerstab entsprechen, zehn Zehnerstäbe einer Hunderterplatte und zehn Hunderterplatten einem Tausenderwürfel
	Erkennen, dass durch Bündelung grosse Anzahlen effizient gezählt werden können
Zahlen darstellen	Stellenwertkarten kennen und nutzen können
	Konventionen der Zahlenschreibweise verstehen und Zahlen schreiben können
	Zahlen in Hunderter, Zehner und Einer zerlegen können
	Erkennen des Zusammenhangs zwischen Dienes-Material und Zifferndarstellung
Stellenwerttafel	Stellenwerttafel kennen und nutzen
	Konventionen der Zahlenschreibweise verstehen
	Erkennen des Zusammenhangs zwischen Material und Zifferndarstellung
	Verschiedene Darstellungen miteinander verknüpfen
Bündeln	Einsicht in das Prinzip der fortgesetzten Bündelung vertiefen
	Abgeschlossene Bündelung direkt in Ziffernschreibweise übersetzen

In der *Unterrichtseinheit „Zahlenstrahl“* wurde der Zahlenraum als lineares Modell thematisiert und die Orientierung im Zahlenraum gefördert. Die Darstellung von Zahlen am Zahlenstrahl wurde erarbeitet. Dabei stand der ordinale Zahlaspekt im Fokus. Die detaillierten Ziele der einzelnen Bausteine der Unterrichtseinheit „Zahlenstrahl“ können der Tabelle 6 entnommen werden.

Tabelle 6: Übersicht über die Ziele der einzelnen Bausteine aus der Unterrichtseinheit „Zahlenstrahl“

Baustein	**Ziele**
Hunderterkette	Zahlenreihe repetieren Struktur der Hunderterkette verstehen Zahlen und Orte finden an der Hunderterkette
Tausenderkette	Analogie der Hunderterketten-Struktur im Tausenderraum erkennen Zahlen und Orte an der Tausenderkette finden
Hunderterstrahl	Zahlen und Orte von der Hunderterkette auf den leeren Zahlenstrahl übertragen
Tausenderstrahl	Am leeren Tausenderstrahl die Hunderterzahlen eintragen Zahlen und Orte auf dem Tausenderstrahl finden Mithilfe des Tausenderstrahls in Schritten zählen
Nachbarzahlen	Sich am leeren Tausenderstrahl durch Streckenhalbieren orientieren Nachbarzahlen am Tausenderstrahl finden Nachbarzehner und Nachbarhunderter am Tausenderstrahl benennen

In der *Unterrichtseinheit „Addition und Subtraktion“* wurden halbschriftliche Rechenstrategien im Tausenderraum erarbeitet und das Bündeln und Entbündeln gefestigt. Zudem wurde das Verständnis für die Mächtigkeit einer Zahl (kardinaler Zahlaspekt) und die Bedeutung der Stellenwerte nochmals vertieft. Die detaillierten Ziele der einzelnen Bausteine der Unterrichtseinheit „Addition und Subtraktion“ können der Tabelle 7 entnommen werden.

Tabelle 7: Übersicht über die Ziele der einzelnen Bausteine aus der Unterrichtseinheit „Addition und Subtraktion“

Baustein	**Ziele**
Zahlen verändern: Bündeln	Mit Dienes-Material und Stellenwertkarten Zahlen verändern Bündeln
Halbschriftliche Addition	Halbschriftliche Addition mit Dienes-Material und Stellenwertkarten darstellen, berechnen und protokollieren Halbschriftliche Addition am Rechenstrich protokollieren
Zahlen verändern: Entbündeln	Mit Dienes-Material und Stellenwertkarten Zahlen verändern Entbündeln
Halbschriftliche Subtraktion: Wegnehmen	Halbschriftliche Subtraktion mit Dienes-Material und Stellenwertkarten darstellen, berechnen und protokollieren Halbschriftliche Subtraktion am Rechenstrich protokollieren
Halbschriftliche Subtraktion: Ergänzen	Halbschriftliche Subtraktion am Rechenstrich durch Ergänzen darstellen, berechnen und protokollieren

Methodische Umsetzung

Der Aufbau der einzelnen Bausteine folgte dem Muster „Einführung, Arbeitsphase, Reflexion und vertiefendes Üben“. Die Einführung und die Reflexion fanden jeweils im Klassenverband statt, um einen Austausch über den gemeinsamen Gegenstand zu ermöglichen. In der *Einführungsphase* erarbeitete die Lehrperson gemeinsam mit den Schülerinnen und Schülern neue

Inhalte, führte diese mit geeigneten Fragestellungen zum Kern der Sache und zeigt Lösungsstrategien und Prozeduren modellhaft vor (vgl. Tabelle 8). Die Einführungssequenzen waren jeweils stark strukturiert.

Tabelle 8: Beispiel einer Einführungssequenz aus dem Baustein „Zahlen darstellen" der Unterrichtseinheit „Dezimalsystem"

„Mit dem Würfelmaterial kann man Zahlen gut darstellen. Ich habe euch dazu heute Zahlenkarten mit Hunderterzahlen (zeigen), Zehnerzahlen (zeigen) und Einerzahlen (zeigen) mitgebracht. Diese Zahlenkarten können euch helfen, herauszufinden, wie unsere Zahlen richtig geschrieben werden."
Die Lehrperson legt die Zahl 254 mit Würfelmaterial und Stellenwertkarten bereit.
„Welche Zahlenkarten passen zum Würfelmaterial? Legt sie dazu. Erklärt, warum die Karten passen."
„Wie viel sind die Platten, Stäbe und Würfel zusammen?" „Wie heisst die Zahl?" „Wie wird die Zahl aufgeschrieben?" usw.

In der *Arbeitsphase* wurden einzeln oder in Partnerarbeit Aufgaben bzw. Arbeitsblätter zum jeweiligen Baustein bearbeitet. Das Ziel der Arbeitsphase war es, ein vertieftes Verständnis für den jeweiligen Inhalt zu ermöglichen bzw. die neu eingeführten Lösungsstrategien anzuwenden. Für die innere Differenzierung lagen für diese Phase jeweils Arbeitsblätter mit differenzierten Aufträgen für unterschiedliche Leistungsniveaus vor (vgl. Tabelle 9).

Tabelle 9: Beispiel einer Arbeitsphase aus dem Baustein „Zahlen darstellen" der Unterrichtseinheit „Dezimalsystem"

Basisniveau	**Erweitertes Niveau**
Zu zweit: Kind A legt eine Anzahl aus dem Tausenderraum mit dem Würfelmaterial. Kind B legt die passenden Stellenwertkarten dazu, setzt die Zahl zusammen und schreibt diese auf eine Blankokarte. Kind A liest die Zahl laut vor. Wechsel. *Die von den Kindern beschrifteten Blankokarten einsammeln und neu an die Kinder verteilen. Sie legen dazu die passenden Stellenwertkarten und das Würfelmaterial.*	*Die Kinder erhalten Stellenwertkarten: zwei Hunderterkarten, zwei Zehnerkarten und zwei Einerkarten:* „Legt mit diesen Karten verschiedene Zahlen, schreibt die Zahlen je auf eine Blankokarte und ordnet sie der Grösse nach. Notiert die Zahlen in dieser Reihenfolge im Rechenheft." *Zu zweit: Ein Kind nennt eine dreistellige Zahl; das andere Kind legt die passende Zahl mit den Stellenwertkarten.*

Die Lehrperson unterstützte in dieser Phase die individuellen Lernprozesse, indem sie spezifische Impulse gab und gezielte Fragen stellte, die dazu anregen sollten vorhandenes Wissen in Worte zu fassen, geeignete Arbeitsmittel zu Hilfe zu nehmen, eigene Aufgaben zu erfinden und über die verwendeten Lösungsstrategien nachzudenken.

In der *Reflexionsphase* wiederum wurden die individuellen Erkenntnisse aus der Arbeitsphase im Klassenverband ausgetauscht und verschiedene Vorgehensweisen und Lösungswege miteinander verglichen und reflektiert. Die Lehrperson half mit geeigneten Fragestellungen und Impulsen, mathematische Muster und Gesetzmässigkeiten zu entdecken und zu beschreiben (vgl. Tabelle 10).

Tabelle 10: Beispiel einer Reflexionsphase aus dem Baustein „Zahlen darstellen“ der Unterrichtseinheit „Dezimalsystem“

Die LP hat im Kreis die Zahlen 736, 420 und 402 mit den Stellenwertkarten gelegt. Die Stellenwertkarten können einzeln hingelegt werden, sodass die Zerlegung der Zahlen sichtbar wird (senkrecht untereinanderlegen).
„Wo ‚sieht‘ man in der geschriebenen Zahl, wie viele Hunderterplatten (Hunderter), Zehnerstäbe (Zehner) und wie viele Einerwürfel (Einer) man braucht, wenn wir die Zahl mit dem Würfelmaterial legen wollen?“
„Was bedeutet die Null in der Zahl 402 (420)?“
„Muss man die Null schreiben, oder kann man sie auch weglassen? Warum?“ usw.

Die Bausteine schlossen mit einer Phase des *vertiefenden Übens*. Ziel dieser Phase war es, Lösungsstrategien einzuschleifen und zu automatisieren. Dafür wurden wiederum Arbeitsblätter für unterschiedliche Leistungsniveaus zur Verfügung gestellt, wobei produktive Übungsformate im Sinne von Wittmann (1992) eingesetzt und Schulbuchseiten sowie Aufgaben aus dem verwendeten Mathematiklehrmittel angegeben wurden, die sich für ein vertieftes Üben eignen.

Zähl- und Kopfrechenkartei

Parallel zu den Unterrichtseinheiten wurde in den ersten sieben Wochen die *Zählkartei* zwei- bis dreimal pro Woche zu Beginn oder am Ende der Lektionen während zehn Minuten eingesetzt. Die Zählkartei lag in drei Schwierigkeitsstufen vor (Zählen in Schritten bis 100, Zählen in Schritten bis 1000 und Zählen in Schritten über 1000). Die Lernziele der Zählkartei können der Tabelle 11 entnommen werden.

Tabelle 11: Übersicht zu den Lernzielen der Zählkartei

Zielgruppe	**Lernziele**
Basisniveau	Sicheres Zählen vorwärts und rückwärts im Hunderterraum in Einer-, Zweier- und Zehnerschritten von beliebigen Startzahlen aus
	Sicheres Zählen vorwärts und rückwärts im Tausenderraum in Einer-, Zweier- und Zehnerschritten
Erweitertes Niveau	Sicheres Zählen vorwärts und rückwärts im Hunderterraum in Fünferschritten von beliebigen Startzahlen aus
	Sicheres Zählen vorwärts und rückwärts im Tausenderraum in Einer-, Zweier-, Zehner- und Hunderterschritten
	Zählen vorwärts und rückwärts über den Tausenderraum hinaus in Einer-, Zweier-, Zehner- und Hunderterschritten

Die *Kopfrechenkartei* wurde ab Woche acht analog zur Zählkartei eingesetzt. Die Kopfrechenkartei war nach dem Prinzip strukturierter Übungsformate nach Wittmann (1992) aufgebaut und umfasste Karteikarten mit verschiedenen Schwierigkeitsgraden zur Addition, zur Subtraktion, zum Verdoppeln, Halbieren und Ergänzen. Sie diente der Automatisierung wichtiger Kernaufgaben. Die Lernziele der Kopfrechenkartei können der Tabelle 12 entnommen werden.

Tabelle 12: Übersicht zu den Lernzielen der Kopfrechenkartei

Zielgruppe	**Lernziele**
Basisniveau	Kennen der Begriffe „Verdoppeln“, „das Doppelte, „Halbieren“ und „die Hälfte“
	Sicheres Verdoppeln und Halbieren bis 20 und das Übertragen auf analoge Aufgaben im Zahlenraum bis 1000
	Sicheres Addieren und Subtrahieren bis 20, das Übertragen auf analoge Rechnungen im Zahlenraum bis 1000 und das Ableiten von ähnlichen Aufgaben
	Sicheres Ergänzen auf 10, auf 100 und auf 1000
Erweitertes Niveau	Kernaufgaben und Beziehungen ausnutzen, um halbschriftliche Aufgaben im erweiterten Zahlenraum zu berechnen

Unterrichtsangebot zum Mathematisieren

Die Förderung enthielt zudem ein Unterrichtsangebot zum Mathematisieren, das alle zwei Wochen als einzelne Lektion oder zwei- bis dreimal pro Woche zu Beginn oder am Ende der Lektion während zehn Minuten eingesetzt werden konnte. Das Unterrichtsangebot bestand aus folgenden Elementen:

- Handlungskarten, bei denen die Lehrperson Rechnungen in Form mathematischer Handlungen am Material dargestellt hat und die Kinder eine dazu passende Rechnung finden mussten,
- Arbeitsblättern mit Rechengeschichten bzw. Bildern, bei denen die dazu passenden Rechnungen gefunden werden mussten,
- Arbeitsblättern, bei denen fehlerhafte Rechnungen, die zu Rechengeschichten bzw. Bildern notiert wurden, identifiziert und korrigiert werden mussten,
- Karten mit Rechengeschichten, Bildern und Rechnungen, die als Domino- oder Memory-Spiel eingesetzt werden konnten.

Dieses Unterrichtsangebot zielte darauf ab, dass mentale Vorstellungsbilder zu den einzelnen Operationen aufgebaut und Verbindungen zwischen Sachsituationen bzw. Bildern und der Gleichungsschreibweise hergestellt wurden (vgl. Tabelle 13).

Tabelle 13: Übersicht zu den Lernzielen des Unterrichtsangebotes zum Operationsverständnis

Inhalt	Lernziele
Aufbau von Grundvorstellungen zu den einzelnen Operationen	Addition: Etwas kommt dazu, oder eine Gesamtmenge wird bestimmt Subtraktion: Etwas geht weg, oder eine Differenz wird bestimmt Multiplikation: Eine Menge wird vervielfacht Division: Eine Menge wird verteilt (Wie viel bekommt jeder?) oder aufgeteilt (Wie viele Gruppen gibt es?)
Verbindung zwischen Sachsituationen/Bildern und der formalen Schreibweise herstellen	Einem Bild eine passende Rechnung zuordnen Einer Handlung eine passende Rechnung zuordnen Zu einer Geschichte eine passende Rechnung finden Zu einer symbolischen Darstellung eine passende Rechnung finden
Sachaufgaben	Einfache Sachaufgaben lösen

Massnahmen und Strategien für eine adaptive Lernbegleitung

Um die Lehrpersonen darin zu unterstützen, eine adaptive Lernbegleitung umzusetzen, enthielten die einzelnen Unterrichtseinheiten jeweils Fragen und Impulse, die die Kinder in ihren Lernprozessen begleiten sollten (vgl. Tabelle 8 und Tabelle 10). Zusätzlich standen den Lehrpersonen Hinweise zur Verfügung, wie sie in den drei Unterrichtsphasen die Schülerinnen und Schüler zum Kern der Sache führen (Einführungsphase), individuelle Unterstützung geben (Arbeitsphase) und den Austausch und Denkprozesse anregen konnten (Reflexionsphase). Einige Beispiele für Impulse und Fragen für die einzelnen Phasen können der Tabelle 14 entnommen werden.

Tabelle 14: Beispiele für Impulse und Fragen zur Unterstützung einer adaptiven Lernbegleitung

Unterrichtsphase	Impulse und Fragen
Einführungsphase: Zum Kern der Sache führen	Was fällt euch auf? Vergleicht! Was ist gleich? Was ist anders? Was bleibt? Wie bzw. wo und woran siehst du das? Kannst du das am Material zeigen? Erkläre mit eigenen Worten, was Kind X meint.
Arbeitsphase: Individuelle Unterstützung geben	Beschreibe, was du gemacht hast! Wo kommst du nicht weiter? Was überlegst du dir? Lege das mit passendem Material! Wie kann man herausfinden, ob das richtig ist?
Reflexionsphase: Austausch und Denkprozesse anregen	Wie habt ihr das herausgefunden? Warum muss man das so machen? Beschreibt die Regeln bzw. das Muster! Hat jemand etwas anderes überlegt? Kann man das noch anders legen bzw. machen, schreiben oder rechnen? Seid ihr einverstanden mit Kind X? Warum? Warum nicht?

5.3.2 Unterschiedlich intensive Form der Begleitung der Lehrpersonen

Die Lehrpersonen der beiden Interventionsgruppen haben an einer dreistündigen Einführungsveranstaltung teilgenommen, in der sie die schriftlichen Anleitungen und die benötigten Materialien erhalten haben. In der Einführungsveranstaltung wurden die Zielsetzung und der zeitliche Ablauf der Studie erläutert, zentrale mathematische Schwierigkeiten rechenschwacher Kinder vermittelt, das Förderkonzept der Intervention vorgestellt und in Beziehung zu den zentralen Lernzielen der dritten Klasse gesetzt. Weiter wurde der Aufbau des Förderprogramms erläutert und die Ziele der einzelnen Förderbestandteile besprochen. Den Lehrpersonen der Gruppe^{Mat} wurde danach Zeit zur Verfügung gestellt, allein oder im Austausch mit anderen Lehrpersonen derselben Interventionsgruppe die schriftlichen Anleitungen und Materialien zu erkunden. Die Lehrpersonen der Gruppe^{Begl} haben hingegen eine zusätzliche Einführungssequenz erhalten, in der die einzelnen Unterrichtseinheiten ausführlich erläutert und detaillierte Hinweise für die Umsetzung gegeben wurden.

Zusätzlich haben die Lehrpersonen der Gruppe^{Begl} während der Interventionszeit an zwei Fortbildungstreffen teilgenommen, die jeweils zweieinhalb Stunden dauerten. Für die Gestaltung dieser Treffen wurde auf Erkenntnisse aus dem Kapitel 3.4 zurückgegriffen. An den Treffen wurden die in den folgenden Wochen umzusetzenden schriftlichen Anleitungen und Materialien der unterrichtsintegrierten Mathematikförderung jeweils ausführlich vorgestellt, zusätzliches Hintergrundwissen vermittelt und die Umsetzung im eigenen Unterricht diskutiert. Zudem wurden Unterrichtsvideos eingesetzt, in denen unterrichtliches Handeln von Lehrpersonen und Lernaktivitäten von Schülerinnen und Schüler zu sehen waren. Diese Videos entstanden im Rahmen der Erprobung der Mathematikförderung. Sie enthielten also Aufnahmen von konkreten Situationen aus der unterrichtsintegrierten Mathematikförderung, zeigten jedoch nicht den eigenen Unterricht der Lehrpersonen, die an den Fortbildungstreffen teilgenommen haben. Die Videos dienten der gemeinsamen Reflexion der Umsetzung der Förderung sowie der gemeinsamen Reflexion der Lern- und Verstehensprozesse der Schülerinnen und Schüler. Es wurden einerseits Videos gezeigt, in denen Lehrpersonen Massnahmen für eine adaptive Lernbegleitung ergriffen haben, andererseits Videos, in denen beobachtet werden konnte, wie Lehrpersonen die Schülerinnen und Schüler zu Erkenntnisprozessen angeregt haben. Zusätzlich boten die Treffen Gelegenheit für einen Austausch über die eigenen Erfahrungen mit der Umsetzung der Förderung sowie für die Beantwortung von Fragen zur Förderung.

5.4 Messinstrumente

In diesem Kapitel werden die Messinstrumente für die Mathematikleistung, die Intelligenz und den SES erläutert, und es wird dargestellt, welche weiteren Daten der Schülerinnen und Schüler bei den Lehrpersonen eingeholt wurden. Alle Testungen, Korrekturen und Dateneingaben wurden von geschulten Testleiterinnen und Testleitern durchgeführt. Die Testleiterinnen und Testleiter haben an einer Einführung in die Testerhebung und Testauswertung teilgenommen. Dazu wurden für alle Testungen detaillierte Anweisungen für Durchführung, Korrektur und Dateneingabe verfasst, sodass eine einheitliche Erhebung und Auswertung gewährleistet werden konnte.

5.4.1 Mathematiktests

Eine zentrale Anforderung an die Mathematiktests bestand darin, dass diese insbesondere auch im Bereich der schwachen mathematischen Leistungen gut zwischen unterschiedlichen Leistungsniveaus differenzieren können. Zudem sollten die Mathematiktests nicht nur Rechenfertigkeiten überprüfen, sondern auch feststellen, ob die Schülerinnen und Schüler zentrale mathematische Inhalte und Konzepte verstanden haben und diese anwenden können. Die Aufgaben sollten daher zentrale Inhalte der Arithmetik abdecken, die für die weitere mathematische Entwicklung von grosser Bedeutung sind und sich auf Inhalte beziehen, die gemäss Curriculum in den Schulen der Deutschschweiz vermittelt werden. Weiter mussten die Tests im Klassenverband durchgeführt werden können. Da keine standardisierten Mathematiktests verfügbar waren, die diese Kriterien erfüllen, wurden Mathematiktests im Rahmen des vorliegenden Projektes teilweise selbst entwickelt und erprobt. Die Instrumente wurden inzwischen veröffentlicht oder stehen kurz vor der Veröffentlichung (Moser Opitz, Stöckli, Grob, Nührenbörger & Reusser, im Druck; Moser Opitz et al., 2016; Moser Opitz, Stöckli, Grob, Nührenbörger & Reusser, 2019).

Pre-Test

Der Mathematiktest der am Ende der zweiten Klasse beim Pre-Test eingesetzt wurde (BASIS-MATH-G 2+; Moser Opitz et al., im Druck), enthält 30 Aufgaben, bei denen maximal 30 Punkte erreicht werden können. Der Test prüft neben den Grundoperationen (Verdoppeln, Halbieren, Ergänzen, Addition, Subtraktion, Multiplikation und Division) insbesondere auch das Verständnis für zentrale mathematische Konzepte (Zahlenstrahl, Bündeln, Zählen, Operationsverständnis, Rechenwege bei Addition, Subtraktion und Textaufgabe, Teile-Ganzes-Verständnis, Mathematisieren; vgl. Tabelle 15) und enthält Aufgaben im Zahlenraum bis hundert. Der Test liegt als Parallelversion in zwei Varianten vor (Testversion A und B). Die Aufgabenreihenfolge wurde z.T. angepasst und die in den Aufgaben verwendeten Zahlen so variiert, dass ein Abschreiben bei einer Durchführung im Klassenverband nicht möglich ist (z.B. Zählen rückwärts: 38, 36, 34, __, __, __, vs. 28, 26, 24, __, __, __,). Die einzelnen Aufgaben können der Tabelle 15 entnommen werden.

Cronbachs α für den Gesamttest beträgt in der vorliegenden Untersuchung .89 (N = 876). Auch die Testhalbierungsreliabilität kann als gut bewertet werden (pro Testteil 15 Aufgaben; erster Testteil: Cronbachs α = .82; zweiter Testteil: Cronbachs α = .82). Für die Testversion A beträgt Cronbachs α = .89, für die Testversion B ist Cronbachs α = .88. Die Trennschärfen dreier Items liegen für die in dieser Arbeit untersuchte Stichprobe zwischen 0.23 und 0.27, diejenigen der anderen Items zwischen 0.31 und 0.61.

Tabelle 15: Aufgaben des Mathematiktests BASIS-MATH-G 2+ (Moser Opitz et al., im Druck) für die Testversion A, eingesetzt beim Pre-Test

Thema	**Aufgabe (Testversion A)**
Zahlenstrahl	
1a)	0 – 100 (17 markieren)
1b)	0 – 100 (86 markieren)
Dezimales Stellenwert-system und Bündeln	
2a)*	47 Punkte, wie viele Zehnerbündel?
2b)*	54 Punkte, wie viele Zehnerbündel?
3a)	3 Schachteln à 10 „Toffifee“, 4 „Toffifee“ einzeln neben den Schachteln, wie viele sind es insgesamt?
3b)	6 Schachteln à 10 „Toffifee“, 2 „Toffifee“ einzeln neben den Schachteln, wie viele sind es insgesamt?
Verdoppeln	
4a)	15
4b)	36
4c)	49
Zählen	
5a)*	38, 36, 34, __, __, __ Zählen in Zweierschritten rückwärts
5b)*	73, 77, 81, __, __, __ Zählen in Viererschritten vorwärts
Ergänzen	
6a)	60 = 15 + __
6b)	76 = 70 + __
6c)	100 = 85 + __
Operationsverständnis	
7a)	Malrechnung mit Ergebnis zum Punktefeld 3 · 7
7b)	Malrechnung mit Ergebnis zum Punktefeld 5 · 4
Addition	
8a)	32 + 47 = __
8b)	Rechenweg zu 9a)

Textaufgabe	
9a)	Lea und Nino haben zusammen 73 Nüsse gesammelt, Lea hat 35 Nüsse. Wie viele Nüsse hat Nino?
9b)	Rechenweg zu 10a)
Zahlen zusammensetzen	
10a)	__ – __ = 11, zwei Zahlen für korrekte Subtraktion auswählen (6, 7, 8, 14, 15, 18), jede Zahl darf in allen Teilaufgaben nur einmal verwendet werden
10b)	__ – __ = 9, zwei Zahlen für korrekte Subtraktion auswählen (6, 7, 8, 14, 15, 18), jede Zahl darf in allen Teilaufgaben nur einmal verwendet werden
10c)	__ – __ = 6, zwei Zahlen für korrekte Subtraktion auswählen· (6, 7, 8, 14, 15, 18), jede Zahl darf in allen Teilaufgaben nur einmal verwendet werden
Sachaufgabe	
11a)	1 Pizza hat 6 Stücke, wie viele Stücke haben 4 Pizzen?
11b)	Rechenweg zu 12a)
Subtraktion	
12a)	73 – 15 = __
12b)	Rechenweg zu 9c)
Halbieren	
13a)	18
13b)	64
13c)	76

* Bei diesen Aufgaben werden in der Testversion B andere Zahlenbeispiele verwendet.

Post-Test und Follow-up 1

Der Mathematiktest für die dritte Klasse (BASIS-MATH-G 3+; Moser Opitz et al., 2019), der beim Post-Test und Follow-up 1 eingesetzt wurde, beinhaltet 41 Aufgaben, in denen 41 Punkte erreicht werden können. Der Mathematiktest enthält Aufgaben im Zahlenraum bis tausend. Auch dieser Test überprüft die Grundoperationen (Verdoppeln, Halbieren, Ergänzen, Addition, Subtraktion, Multiplikation und Division) und das Verständnis für zentrale mathematische Konzepte (Zahlenstrahl, Bündeln und Entbündeln, Zählen, Operationsverständnis, Rechenwege bei Addition, Subtraktion, Text- und Sachaufgaben sowie Mathematisieren; vgl. Tabelle 16). Da dieser Test sowohl beim Post-Test als auch beim Follow-up 1 zum Einsatz kam, haben die Schülerinnen und Schüler, die beim Post-Test die Testversion A bearbeitet haben, beim Follow-up 1 die Testversion B gelöst und umgekehrt. Damit können Leistungssteigerungen, die durch das erneute Bearbeiten der genau gleichen Aufgaben bedingt sein können, gemindert werden. Die einzelnen Aufgaben können für die Testversion A der Tabelle 16 entnommen werden.

Beim Post-Test ergibt sich in der vorliegenden Untersuchung (n = 870) eine sehr hohe interne Konsistenz des Mathematiktests für die dritte Klasse von Cronbachs α = .92. Die Testhalbierungsreliabilität beträgt Cronbachs α = .88 beim ersten Testteil (21 Aufgaben) und Cronbachs α

= .84 beim zweiten Testteil (20 Aufgaben). Für die Testversion A ist Cronbachs α = .93, für die Testversion B ist Cronbachs α =.92. Die Itemanalyse zeigt, dass mit Ausnahme einer Aufgabe (r_{it} = .29, Aufgabe 10a; vgl. Tabelle 16) alle Trennschärfen über dem als kritisch zu bewertenden Wert von r_{it} = .30 liegen (niedrigster Wert: r_{it} = .34; höchster Wert r_{it} = .64).

Für den Follow-up 1 (N = 877) beträgt Cronbachs α = .92, und auch die Testhalbierungsreliabilität weist gute Werte auf (erster Testteil mit 21 Aufgaben: Cronbachs α = .89; zweiter Testteil mit 20 Aufgaben: Cronbachs α = .83). Für die Testversion A ist Cronbachs α = .92, für die Testversion B ist Cronbachs α = .93. Die Trennschärfe dreier Items liegt zwischen 0.22 und 0.27, diejenigen der anderen Items zwischen 0.34 und 0.64. Da der Mathematiktest für die dritte Klasse zweimal eingesetzt wurde, kann die Retest-Reliabilität berechnet werden, die bei r = .88 liegt ($p < .001$).

Tabelle 16: Aufgaben des Mathematiktests BASIS-MATH-G 3+ (Moser Opitz et al., 2019) für die Testversion A, eingesetzt beim Post-Test und beim Follow-up 1

Thema	**Aufgabe (Testversion A)**
Zahlenstrahl	
1a)	1 – 1000 (860 markieren)
1b)	1 – 1000 (473 markieren)
Zählen	
2a)*	477, 482, 487, __, __, __ Zählen in Fünferschritten vorwärts
2b)*	635, 625, 615, __, __, __ Zählen in Zehnerschritten rückwärts
Halbieren	
3a)	76
3b)	124
3c)	360
Ergänzen	
4a)	37 + __ = 50
4b)	125 + __ = 300
Subtraktion	
5a)	73 – 15 = __
5b)	83 – 34 = __
5c)	95 – 59 = __
Entbündeln	
6a)	1000 – 10 = __
6b)	1000 – 100 = __
6c)	1000 – 110 = __
Sachaufgabe	
7a)*	Andi kauft 1 kg Birnen, 2 kg Äpfel, 3 Zitronen. Wie viel muss er bezahlen? (Kilopreis Äpfel: 3.70 Fr., Kilopreis Birnen: 2.60 Fr., Preis pro Zitrone: 80 Rp.)
7b)*	Rechenweg zu 7a)

Verdoppeln	
8a)	36
8b)	209
8c)	317
Geld zählen	
9)*	Wie viel Geld ist in der Klassenkasse? Tabellarische Darstellung (Anzahl Noten / Münzen): 1 mal 100 Fr., 4 mal 10 Fr., 13 mal 1 Fr.
Bündeln	
10a)*	8 Hunderter, 27 Einer. Notiere die Zahl.
10b)*	3 Hunderter, 13 Zehner, 6 Einer. Notiere die Zahl.
Sachaufgabe	
11a)	1 Brötchen kostet 50 Rp. Wie viel kosten 5 Brötchen?
11b)	Wie viel kosten 15 Brötchen?
12a)*	Lea hat 50 Nüsse, Reto hat 20 Nüsse mehr als Lea, und Stefan hat 30 Nüsse weniger als Reto. Wie viele Nüsse haben alle zusammen?
12b)*	Rechenweg zu 12a)
Operationsverständnis	
13a)*	Malrechnung mit Ergebnis zum Punktefeld 3 · 8
13b)	Malrechnung mit Ergebnis zum Punktefeld 4 · 5
Multiplikation	
14a)	10 · 40 = __
14b)	5 · 30 = __
Division	
15a)	100 : 25 = __
15b)	1000 : 5 = __
Operationsverständnis	
16a)*	3 · 6 = 18, Multiplikation in Division umwandeln
16b)*	35 : 7 = 5, Division in Multiplikation umwandeln
Addition	
17a)*	457 + 246 = __
17b)*	Rechenweg zu 17a)
Subtraktion	
18a)*	365 – 128 = __
18b)*	Rechenweg zu 18a)
Sachaufgabe	
19a)*	Der Vater kauft das Velo und den Helm. Er gibt 1000 Fr. Was bekommt er zurück? (Preis für Velo: 599.– und Preis für Helm: 129.–)
19b)*	Rechenweg zu 19a)

* Bei diesen Aufgaben werden in der Testversion B andere Zahlenbeispiele verwendet.

Follow-up 2

Der Mathematiktest für Ende vierter Klasse, der beim Follow-up 2 eingesetzt wurde (BASIS-MATH-G 4^{+}–5; Moser Opitz et al., 2016), enthält 58 Aufgaben, bei denen insgesamt 60 Punkte erreicht werden können.[13] Es wurde die Schweizer Testversion für das vierte Quartal der vierten Klasse verwendet. Der Mathematiktest enthält Aufgaben im Zahlenraum bis zehntausend. Der Test erfasst basale Kompetenzen der Grundschulmathematik in den Bereichen Arithmetik und Sachrechnen und überprüft den Umgang mit Zahlen, Massen, Grundoperationen und Rechenverfahren. Eine Auswahl von Aufgaben ist in der Tabelle 17 für die Testversion A aufgeführt, für die restlichen Aufgaben wird auf den veröffentlichten Test verwiesen.

In der vorliegenden Stichprobe (N = 835) beträgt die interne Konsistenz Cronbachs α = .92. Dieser Wert ist identisch mit dem Wert, der für die Normierungsstichprobe der Schweiz im Testhandbuch angegeben wird (Moser Opitz et al., 2016, S. 49). Für die Testhalbierungsreliabilität liegen die Werte in dieser Untersuchung beim ersten Testteil bei Cronbachs α = .85 und beim zweiten Testteil bei Cronbachs α = .88 (pro Testteil 29 Aufgaben). Für die Testvarianten A und B beträgt Cronbachs α = .92 für beide Versionen. Gemäss Testhandbuch beträgt die mittlere Trennschärfe in der gesamten Normierungsstichprobe r_{it} = .46, wobei nur für zwei Aufgaben eine Trennschärfe kleiner als r_{it} = .30 vorliegt. Die Trennschärfen in der Gesamtstichprobe der vorliegenden Untersuchung (N = 835) liegen bei 13 Aufgaben zwischen 0.22 und 0.29, diejenigen der restlichen Items zwischen 0.30 und 0.61. Bei den Aufgaben mit niedrigen Trennschärfen handelt es sich jedoch fast ausschliesslich um Aufgaben, die das untere Leistungsspektrum abdecken und somit wichtig sind, um auch im unteren Leistungsbereich ausreichend differenzieren zu können.

Tabelle 17: Eine Auswahl von Aufgaben des Mathematiktests BASIS-MATH-G 4^{+}–5 (Moser Opitz et al., 2016), eingesetzt beim Follow-up 2

Thema	**Aufgabe (Testversion A)**
Kopfrechnen	
1a)*	27 + 48 = ___
1d)*	1 Fr. 95 Rp. + 3 Fr. 55 Rp. = ___
1e)*	3 Fr. 80 Rp. – 1 Fr. 65 Rp. = ___
Zahlen zerlegen	
2a)*	75 = ___ + ___
2b)*	75 = ___ + 1 + ___
2c)*	75 = ___ + ___ + ___ + ___
Halbieren	
5d)	24 : 2 = ___
5e)	280 : 2 = ___
5f)	360 : 2 = ___

[13] Zwei Aufgaben sind Partial-Credit-Items, bei denen je nach Anzahl korrekt angekreuzter Antworten ein oder zwei Punkte vergeben werden; 18a) Grössen: Gewicht, 18b) Grössen: Längen; vgl. Tabelle 17.

Textaufgabe	
12a)	Welche Rechnung passt zur Textaufgabe? Ein Schäferhund wiegt 42 kg. Ein Dalmatiner ist um 7 kg leichter. Wie schwer ist der Dalmatiner? 7 · 6, 42 : 7, 42 – 7, 42 × 7, 42 + 7.
Sachaufgabe	
18a)[1)]	Drei Taschen sind genau gleich schwer. Welche? Kreuze die drei gleich schweren Taschen an. 4,3 kg, 430 g, 4300 g, 4 kg 300 g, 43 g

* Bei diesen Aufgaben werden in der Testversion B andere Zahlenbeispiele verwendet; [1)] Wenn alle drei angekreuzten Antworten korrekt sind, werden zwei Punkte vergeben. Werden nur zwei korrekte Antworten angekreuzt, wird ein Punkt vergeben.

5.4.2 Intelligenztest

Für die Erhebung des IQs wurde der CFT 1 (Cattell, Weiß & Osterland, 1997) verwendet. Es handelt sich hierbei um einen Intelligenztest, der für sich in Anspruch nimmt, die Intelligenz nahezu unabhängig von der sozialen bzw. der regionalen Herkunft sowie weitgehend unbeeinflusst von bestehenden Lernerfahrungen insbesondere im sprachlichen Bereich zu erfassen (Weiß & Osterland, 1997). Der Test baut auf der Intelligenztheorie von Cattell auf (z.B. Cattell, 1965; vgl. Weiß & Osterland, 1997, siehe auch Kap. 2.3.2) und erfasst primär die fluide Intelligenz, hingegen weniger stark die kristalline Intelligenz (Weiß & Osterland, 1997). Der Test besteht aus fünf Untertests (Substitutionen, Labyrinthe, Klassifikation, Ähnlichkeiten und Matrizen), die unter verschiedenen Zeitvorgaben gelöst werden müssen (Weiß & Osterland, 1997). Der Test kann als Gruppentest durchgeführt werden und liegt in zwei Parallelformen A und B vor, sodass durch Änderungen in der Reihenfolge die Möglichkeit zum Abschreiben minimiert wird (Weiß & Osterland, 1997).

Für den CFT 1 (Cattell et al., 1997) liegen Normen ab dem Alter von 7 Jahren und 6 Monaten (90 Monate) bis zu 9 Jahren und 5 Monaten (113 Monate) vor. Bei Kindern aus der vorliegenden Stichprobe, die die höchste Altersnorm überschritten haben, wurde eine annähernde Schätzung des IQs vorgenommen, indem auf die höchste vorhandene Altersnorm zurückgegriffen wurde ($n = 45$; vgl. Kapitel 5.5). Das führt dazu, dass die IQ-Werte für diese Kinder tendenziell eher zu hoch eingeschätzt sind. Die daraus entstehenden Implikationen für die Analysen und die Ergebnisse können den Kapiteln 5.5 und 5.6 entnommen werden.

Die interne Konsistenz für die verschiedenen Untertests liegt gemäss Testhandbuch für beide Testversionen in der zweiten Klasse zwischen $r = .65$ und $r = .86$ und für die Summe der Untertests für die Testversion A bei $r = .92$ bzw. für die Testversion B bei $r = .90$ (Weiß & Osterland, 1997, S. 24). Zudem liegen Angaben zur Kriteriumsvalidität vor, wonach die Korrelation zwischen dem CFT 1 (Cattell et al., 1997) und dem HAWIK-Intelligenztest (Hardesty & Priester, 1966) r = .66 beträgt (Cattell et al., 1997, S. 31). Der Test erfüllt somit gängige Anforderungen an die Testgütekriterien.[14]

5.4.3 Sozioökonomischer Status

Der sozioökonomische Status (SES) wurde in der vorliegenden Arbeit mithilfe der *Bücheraufgabe* von Paulus (2009) erfasst. Die Bücheraufgabe besteht aus einer bildlichen Darstellung des familiären Buchbestandes sowie einer verbalen Beschreibung des jeweiligen Buchbestandes mit fünf Antwortmöglichkeiten (Paulus, 2009, vgl. Abbildung 9). Die Bücheraufgabe baut auf dem Begriff des kulturellen Kapitals von Bourdieu (1983) auf, wobei die Bücheraufgabe das kulturelle Kapital im Sinne des „objektivierten Zustands" erfassen soll (Paulus, 2009).

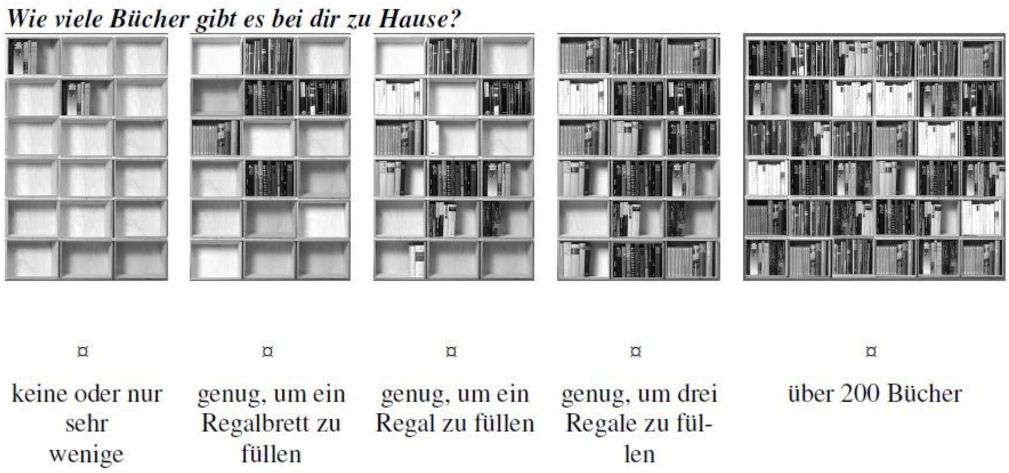


Abbildung 9: Bücheraufgabe gemäss Paulus (2009, S. 5)[15]

Bei der Bücheraufgabe handelt es sich um eine sehr ökonomische Erfassungsmöglichkeit des sozioökonomischen Status, da keine Befragung der Eltern notwendig ist und sich in verschiedenen Studien gezeigt hat, dass der familiäre Buchbestand ein guter Indikator für den sozioökonomischen Status einer Familie ist (Paulus, 2009). Es wird empfohlen, die Bücheraufgabe nach Möglichkeit zweimal einzusetzen, da insbesondere bei jüngeren Kindern beim zweiten Mal von einer realistischeren Schätzung auszugehen ist (Paulus, 2009). In der vorliegenden Arbeit wurde die Bücheraufgabe daher sowohl beim Pre-Test am Ende der zweiten Klasse als auch beim Follow-up 2 am Ende der vierten Klasse erfasst. Für die statistischen Analysen wurde der SES gemäss der zweiten Erhebung der Bücheraufgabe beim Follow-up 2 verwendet ausser bei Kindern, bei denen nur bei der ersten Erhebung der SES dokumentiert werden konnte.

[14] Eine Berechnung der Testgütekriterien auf Basis der Gesamtstichprobe der vorliegenden Arbeit war nicht möglich, da die einzelnen Items des Tests nicht separat digitalisiert wurden, sondern nur der Gesamtwert des IQs im Datensatz erfasst wurde.

[15] Da im deutschschweizerischen Kontext die Begriffe Regalbrett und Regal eher selten verwendet werden, wurden für die Erfassung des SES in der vorliegenden Arbeit diese Begriffe durch die Begriffe Tablar und Gestell ersetzt.

Dabei wurde eine Skalierung von eins bis fünf gewählt, wobei der Wert 1 für keine oder nur sehr wenige Bücher steht. Gemäss einer Untersuchung von Paulus (2009, S. 6), liegt die Retest-Reliabilität bei $r = .81$. In der vorliegenden Arbeit beträgt die Retest-Reliabilität $r = .57$ ($n = 826$, $p < .01$). Dass dieser Wert niedriger ausfällt, kann darauf zurückgeführt werden, dass bei Paulus (2009) der zeitliche Abstand zwischen den beiden Messungen mit sechs Tagen sehr kurz war.

5.4.4 Weitere erhobene Daten

Zusätzlich zu den bisher genannten Messinstrumenten wurden beim Pre-Test weitere Daten zu den Schülerinnen und Schülern bei den Lehrpersonen eingeholt:

- Geburtsdatum
- Geschlecht
- Förderbedarf in Mathematik
- Sprachstand: Deutschverstehen

Beim Sprachstand wurden die Lehrpersonen auf einer vierstufigen Skala nach ihrer Einschätzung gefragt, wie gut das Kind Deutsch versteht (keine Einschränkungen, geringe Einschränkungen, erhebliche Einschränkungen, gar nicht oder fast gar nicht). Des Weiteren wurden die Lehrpersonen gebeten, Kinder mit einem besonderen Förderbedarf in Mathematik zu benennen. Dabei handelt es sich um Kinder mit unterdurchschnittlichen Leistungen im Fach Mathematik, die gemäss Einschätzung der Lehrpersonen auf zusätzliche Unterstützung im Mathematikunterricht angewiesen sind. Es kann jedoch nicht gesagt werden, inwiefern diese Einschätzung allenfalls auf einem von einer Fachperson mit einem standardisierten Test diagnostizierten Förderbedarf beruht. Sowohl beim Förderbedarf als auch beim Sprachstand muss daher beachtet werden, dass es sich um eine Einschätzung durch die Lehrpersonen handelt, die durch das allgemeine Leistungsniveau der aktuellen oder der früheren Klassen beeinflusst sein kann. Die Einschätzung des Sprachstandes wurde für die Analyse als Dummy-Variable codiert (0 = keine Einschränkungen, 1 = geringe Einschränkungen, erhebliche Einschränkungen, gar nicht oder fast gar nicht).

5.4.5 Klassenvariablen

Bei Untersuchungen im Klassenkontext stellen klassenbezogene Merkmale eine wichtige Ergänzung zu Individualmerkmalen dar (Brühwiler et al., 2017). Da in der vorliegenden Untersuchung zudem statistische Analyseverfahren eingesetzt werden (vgl. Kap. 5.6.2), die sowohl Variablen auf Individual- als auch auf Klassenebene berücksichtigen können, wurden aus den bisher genannten Individualmerkmalen auch klassenweise aggregierte Variablen gebildet. Daher können folgende Variablen nicht nur auf Individual-, sondern auch auf Klassenebene für die statistischen Analysen verwendet werden:

- Mathematisches Vorwissen der Klasse (Mathematikleistung Pre-Test)
- IQ-Mittelwert der Klasse
- SES-Mittelwert der Klasse
- Durchschnittliches Alter der Klasse
- Anteil Knaben in den Klassen (%)
- Anteil Kinder mit Förderbedarf in Mathematik (%)
- Anteil Kinder mit Einschränkungen im Deutschverstehen (%)

5.5 Deskriptive Beschreibung der Gesamtstichprobe und der Untersuchungsstichproben für die Hypothesenprüfung

Im Folgenden wird zunächst die Gesamtstichprobe der vorliegenden Untersuchung beschrieben, es wird aufgezeigt, wie die erfassten Merkmale über die drei Stichprobengruppen verteilt sind, und es wird geklärt, ob es zwischen den Stichprobengruppen signifikante Unterschiede beim Pre-Test gibt. Im Anschluss daran werden die Untersuchungsstichproben, die für die Beantwortung der Fragestellungen verwendet wurden, näher erläutert.

5.5.1 Gesamtstichprobe

Die Gesamtstichprobe bestand aus 61 Schulklassen. Drei Klassen der Interventionsgruppe^{Mat} konnten jedoch nicht in die Auswertungen einbezogen werden, da in einem Fall ein Ausstieg aus dem Projekt gewünscht wurde und in zwei Fällen die Intervention nicht gemäss den Vorgaben durchgeführt wurde.[16] Die Gesamtstichprobe für die Analysen setzt sich daher aus 58 Schulklassen aus neun Deutschschweizer Kantonen (Aargau, Basel-Landschaft, Basel-Stadt, Bern, Luzern, Solothurn, Thurgau, Zug und Zürich) zusammen und umfasst insgesamt 888 Schülerinnen und Schüler, die während der Interventionsphase die dritte Klasse besucht haben.[17] In den genannten Kantonen werden Schülerinnen und Schüler mit Rechenschwäche und anderen Schulleistungsschwierigkeiten mehrheitlich innerhalb von Regelklassen beschult. Auf-

[16] In einer Klasse hat nur die Schulische Heilpädagogin mit Kindern mit einem integrativen Förderbedarf separiert von der restlichen Klasse die Förderung durchgeführt. In der anderen Klasse hat die Lehrperson nur ausgewählte Bestandteile der Förderung mit der ganzen Klasse umgesetzt.

[17] Zusätzlich haben 23 Schülerinnen und Schüler beim ersten Messzeitpunkt Ende zweiter Klasse an den Erhebungen teilgenommen. Da diese auf das neue Schuljahr hin das Schulhaus gewechselt oder eine Klasse repetiert haben, haben sie nicht an der Intervention teilgenommen und wurden für die vorliegenden Analysen weggelassen.

grund des föderalistisch aufgebauten Bildungssystems in der Schweiz können sich die Rahmenbedingungen für die Beschulung dieser Kinder in Regelklassen jedoch teilweise unterscheiden. Im Grundsatz kann aber festgehalten werden, dass Schülerinnen und Schüler, bei denen aufgrund von Lernschwierigkeiten im Fach Mathematik von einem speziellen Förderbedarf ausgegangen wird, im Rahmen von Integrativer Förderung (IF) zusätzliche Unterstützung erhalten. Dabei ist vorgesehen, dass diese Förderung durch entsprechend ausgebildete Schulische Heilpädagoginnen und Heilpädagogen erfolgt. Die personellen und zeitlichen Ressourcen variieren diesbezüglich stark, sodass auch keine generalisierenden Aussagen zu den jeweiligen Umsetzungsformen möglich sind (z.B. Einzelförderung, Kleingruppenförderung, Förderung innerhalb des Klassenunterrichts etc.).

In der Stichprobe der vorliegenden Untersuchung sind sowohl Schulen aus Städten, der Agglomeration als auch aus ländlichen Gebieten vertreten. Die durchschnittliche Klassengrösse liegt bei 15.31 (min. = 5, max. = 24, *SD* = 5.24). Dabei muss beachtet werden, dass nur diejenigen Kinder, für die eine Einverständniserklärung der Erziehungsberechtigten für die Teilnahme an der Untersuchung vorlag, für diese Berechnung berücksichtigt werden konnten, sodass die tatsächliche Anzahl der Kinder pro Klasse von den genannten Werten teilweise (stark) abweichen kann. Bei zehn Klassen handelt es sich zudem um altersdurchmischte Klassen, bei denen nur die Kinder der dritten Klasse an den Erhebungen teilgenommen haben (vgl. Tabelle 19).[18]

Für die 888 Schülerinnen und Schüler liegen nicht für alle Variablen und nicht für alle Messzeitpunkte Daten vor (vgl. Tabelle 18). Es gibt zwei unterschiedliche Gründe für fehlende Daten. Einerseits waren gewisse Schülerinnen und Schüler bei einem oder mehreren Erhebungszeitpunkten krank, und eine Nachtestung konnte nicht organisiert werden. Andererseits haben einzelne Kinder den Wohnort und somit das Schulhaus gewechselt oder mussten eine Klasse repetieren. Da mit Ausnahme vereinzelter Repetitionen keine systematischen Verzerrungen durch die fehlenden Werte zu erwarten sind und die Anzahl fehlender Werte im Vergleich zur Gesamtstichprobengrösse als eher klein einzuschätzen ist, wurde auf eine Imputation der fehlenden Daten verzichtet (vgl. Tabelle 18). Zusätzlich muss beim IQ berücksichtigt werden, dass insgesamt bei 871 Kindern eine IQ-Messung vorliegt (vgl. IQ „alle“ in Tabelle 18). Bei 45 Kindern konnte jedoch der IQ-Wert nur auf Basis der höchsten vorhandenen Altersnorm berechnet werden, da diese Kinder beim Erhebungszeitpunkt ein höheres Alter aufwiesen, als in der Altersnorm des Tests vorgesehen ist (vgl. Kap. 5.4.2). Daher wird separat auch berichtet,

[18] Davon acht Klassen, in denen Kinder der dritten und der vierten Klasse gemeinsam eine Klasse besuchten, und zwei Klassen mit Kindern der ersten bis dritten Klasse.

für wie viele Kinder IQ-Werte vorliegen, bei denen eine reguläre IQ-Bestimmung auf Basis der offiziellen Altersnorm möglich war (vgl. IQ „streng“ in Tabelle 18).

Tabelle 18: Übersicht über die Anzahl vorhandener Daten (*n*) in der Gesamtstichprobe

Variable	**Anzahl vorhandener Daten (*n*)**
Vorwissen Mathematik (Pre-Test)	876
Mathematik Post-Test	870
Mathematik Follow-up 1	877
Mathematik Follow-up 2	835
Alter	878
IQ „alle“ (alle Kinder, inkl. der Kinder, bei denen der IQ auf Basis der höchsten vorhandenen Altersnorm bestimmt wurde)	871
IQ „streng“ (nur Kinder, deren Alter innerhalb der offiziellen Altersnorm lag)	826
SES	887
Geschlecht	888
Förderbedarf Mathematik	888
Einschränkungen im Deutschverstehen	886

Bei der randomisierten Zuteilung der Klassen auf die drei Stichprobengruppen wurde auf eine gleichmässige Verteilung der Schulen aus den verschiedenen Einzugsgebieten (Stadt, Land, Agglomeration) sowie hinsichtlich des sozioökonomischen Hintergrunds geachtet. Klassen aus dem gleichen Schulhaus wurden zudem in die gleiche Stichprobengruppe eingeteilt, um kontaktbedingte Verzerrungen zu vermeiden. Eine gleichmässige Verteilung der altersdurchmischten Klassen auf die drei Stichprobengruppen konnte jedoch nicht berücksichtigt werden. In der Interventionsgruppe^{Begl} waren 21 Klassen mit 281 Schülerinnen und Schülern eingeteilt, in der Interventionsgruppe^{Mat} 16 Klassen mit 288 Schülerinnen und Schülern und in der Kontrollgruppe 21 Klassen mit 319 Schülerinnen und Schülern (vgl. Tabelle 19).

Tabelle 19: Stichprobengrösse (Anzahl Klassen und Anzahl Kinder) für die Gesamtstichprobe und die drei Stichprobengruppen

	Gesamtstichprobe	**Gruppe^{Begl}**	**Gruppe^{Mat}**	**Gruppe^{Kontr}**
Anzahl Klassen	58	21	16	21
Anzahl Kinder	888	281	288	319
Anzahl altersdurchmischter Klassen	10	7	1	2
Anzahl Kinder in altersdurchmischten Klassen	81	50	10	21

Die Verteilung zentraler Merkmale in der Gesamtstichprobe und den drei Stichprobengruppen kann der Tabelle 20 und der Tabelle 21 entnommen werden. Die Tabelle 20 enthält die deskriptiven Statistiken zu dem mathematischen Vorwissen, dem Alter, dem IQ und dem SES beim

Pre-Test unter Berücksichtigung aller vorhandenen Daten. Beim IQ wird sowohl der Mittelwert für alle Kinder, für die eine IQ-Messung vorliegt, berichtet (IQ „alle“) als auch der Mittelwert für die Kinder, deren Alter innerhalb der offiziellen Altersnorm liegt (IQ „streng“; vgl. Tabelle 20). Die Tabelle 21 informiert über die Häufigkeitsverteilung nominaler Merkmale (Geschlecht, Förderbedarf in Mathematik, Einschränkungen im Deutschverstehen) unter Berücksichtigung aller vorhandenen Daten.

Tabelle 20: Verteilung des Vorwissens in Mathematik, des Alters, des IQs und des SES in der Gesamtstichprobe und den drei Stichprobengruppen unter Berücksichtigung aller vorhandenen Daten

	Gesamtstichprobe	**Gruppe^{Begl}**	**Gruppe^{Mat}**	**Gruppe^{Kontr}**
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.15 (6.36)	20.86 (6.57)	20.56 (6.53)	21.94 (5.93)
Alter (in Monaten)	103.97 (5.53)	103.84 (5.97)	104.85 (5.42)	103.29 (5.13)
IQ „alle“	105.13 (14.06)	104.03 (15.57)	104.27 (13.53)	106.87 (12.96)
IQ „streng“	105.62 (14.03)	104.53 (15.69)	104.82 (13.38)	107.24 (12.97)
SES	3.84 (1.16)	3.74 (1.21)	3.76 (1.23)	4.00 (1.04)

Für die Gesamtstichprobe beträgt der Mittelwert für das mathematische Vorwissen 21.15 Punkte (SD = 6.36; vgl. Tabelle 20). Trotz randomisierter Zuteilung der Klassen auf die drei Stichprobengruppen zeigt sich, dass sich die Mittelwerte für das mathematische Vorwissen in den drei Stichprobengruppen signifikant unterscheiden ($F[2,873] = 3.97$, $p = .019$, partielles $\eta^2 = .009$). Post-hoc-Tests belegen, dass die Gruppe^{Mat} über signifikant weniger mathematisches Vorwissen verfügt als die Gruppe^{Kontr} ($p = .022$). Die Gesamtstichprobe ist beim Pre-Test im Durchschnitt 103.97 Monate alt (SD = 5.53). Auch hier liegen signifikante Unterschiede zwischen den Stichprobengruppen vor ($F[2,875] = 6.16$, $p = .002$, partielles $\eta^2 = .014$). Die Schülerinnen und Schüler der Gruppe^{Mat} sind signifikant älter als diejenigen der Gruppe^{Kontr} ($p = .002$) und tendenziell älter als diejenigen der Gruppe^{Begl} ($p = .073$). Der Mittelwert für den IQ liegt in der Gesamtstichprobe für die IQ-Variable „alle“ bei 105.13 Punkten (SD = 14.06) und für die IQ-Variable „streng“ bei 105.62 Punkten (SD = 14.03). Für beide IQ-Variablen müssen abweichende Mittelwerte für die Stichprobengruppen angenommen werden (IQ „alle“: Welch-Test: $F[2,565] = 4.012$, $p = .019$, partielles $\eta^2 = .009$; IQ „streng“: Welch-Test: $F[2,532] = 3.413$, $p = .034$, partielles $\eta^2 = .008$). Für die IQ-Variable „streng“ kann jedoch trotz festgestellter Mittelwertsunterschiede in der einfaktoriellen Varianzanalyse nur ein tendenzieller Unterschied der Gruppe^{Begl} ($p = .072$) sowie der Gruppe^{Mat} ($p = .075$) im Vergleich zur Gruppe^{Kontr} beobachtet werden. Für die IQ-Variable „alle“ hingegen zeigen sich signifikante Unterschiede zwischen der Gruppe^{Kontr} und den beiden Interventionsgruppen (Gruppe^{Begl}: $p = .047$; Gruppe^{Mat}: $p =$

.046), wobei für die Gruppe^{Kontr} höhere IQ-Werte vorliegen. Beim SES beträgt der Mittelwert für die Gesamtstichprobe 3.84 Punkte (SD = 1.16), wobei auch hier für die Stichprobengruppen signifikant unterschiedliche Mittelwerte angenommen werden müssen (Welch-Test: $F[2,574] = 5.081$, $p = 0.06$, partielles $\eta^2 = .010$). Gemäss Post-hoc-Tests weist die Gruppe^{Kontr} einen signifikant höheren SES auf als die Gruppe^{Begl} ($p = .021$) und die Gruppe^{Mat} ($p = .030$).

Für die nominal skalierten Daten zeigt sich unter Berücksichtigung aller vorhandenen Daten in der Gesamtstichprobe eine ausgeglichene Geschlechterverteilung (50%; vgl. Tabelle 21). Es liegen keine signifikanten Unterschiede zwischen den Stichprobengruppen vor ($\chi^2[2, N = 888] = 4.631$, $p = .099$, Cramers V = .072, $p = .099$). Gemäss Angaben der Lehrpersonen weisen in der Gesamtstichprobe 20.3% der Kinder einen Förderbedarf in Mathematik auf, für die Gruppe^{Mat} liegt dieser Wert deutlich höher (26.4%). Die Unterschiede in der Häufigkeit sind zwischen den Stichprobengruppen signifikant ($\chi^2[2, N = 888] = 10.092$, $p = .006$, Cramers V = .107, $p = .006$), wobei die standardisierten Residuen zeigen, dass in der Gruppe^{Mat} mehr Kinder einen Förderbedarf in Mathematik haben ($z = 2.3$). Der Anteil Kinder mit Einschränkungen im Deutschverstehen liegt gemäss Lehrpersonen in der Gesamtstichprobe bei 23.9%. Die Unterschiede in den Häufigkeiten zwischen den Stichprobengruppen sind hier knapp nicht signifikant ($\chi^2[2, N = 886] = 5.735$, $p = .057$, Cramers V = .080, $p = .057$).

Tabelle 21: Verteilung des Geschlechts, des Förderbedarfs in Mathematik und der Einschränkungen im Deutschverstehen in der Gesamtstichprobe und den drei Stichprobengruppen unter Berücksichtigung aller vorhandenen Daten

	Gesamtstichprobe	**Gruppe^{Begl}**	**Gruppe^{Mat}**	**Gruppe^{Kontr}**
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	444 (50.0)	148 (52.7)	129 (44.8)	167 (52.4)
Knaben	444 (50.0)	133 (47.3)	159 (55.2)	152 (47.6)
Förderbedarf Mathematik	180 (20.3)	51 (18.1)	76 (26.4)	53 (16.6)
Kinder mit Einschränkungen im Deutschverstehen	212 (23.9)	81 (28.8)	64 (22.2)	67 (21.0)

Mit Blick auf die Effektstärken kann festgehalten werden, dass es sich bei allen oben genannten signifikanten Unterschieden zwischen den Stichprobengruppen um eher kleine Effekte handelt, da bei keiner Variablen der Schwellenwert für mittlere Effekte von .5 erreicht wird (Cohen, 1988). Da sich solche Gruppenunterschiede in Pre-Test-Merkmalen dennoch auf die Ergebnisse von quasi-experimentellen Untersuchungen auswirken können, ist es unabdingbar, dass in den zu berechnenden statistischen Modellen die entsprechenden Variablen als Kontrollvariablen einbezogen werden. Nur so kann sichergestellt werden, dass allfällig festzustellende Interventionseffekte tatsächlich auf die Intervention und nicht auf die unterschiedlichen Gruppenvoraussetzungen zurückzuführen sind.

Die Verteilung der abhängigen Variablen in der Gesamtstichprobe und den drei Stichprobengruppen kann der Tabelle 22 entnommen werden. Beim Post-Test direkt im Anschluss an die Intervention liegt der Mittelwert für die Gesamtstichprobe bei 25.49 Punkten (*SD* = 9.03), beim Follow-up 1 bei 27.62 Punkten (*SD* = 8.74) und beim Follow-up 2 bei 44.01 Punkten (*SD* = 10.50). An dieser Stelle soll nochmals daran erinnert werden, dass beim Post-Test und beim Follow-up 1 der gleiche Mathematiktest verwendet wurde, sodass die Punktezahl für diese beiden Messzeitpunkte direkt miteinander verglichen werden kann. Ohne Kontrolle der Pre-Test-Unterschiede zwischen den Stichprobengruppen und ohne Berücksichtigung der hierarchischen Datenstruktur zeigt sich, dass die $\text{Gruppe}^{\text{Begl}}$ die niedrigsten Mittelwerte bei allen drei Messzeitpunkten nach der Intervention aufweist und in der $\text{Gruppe}^{\text{Kontr}}$ bei allen drei Messzeitpunkten die höchsten Werte erreicht werden. Die Mittelwerte für die $\text{Gruppe}^{\text{Mat}}$ liegen zwischen den Mittelwerten der beiden anderen Stichprobengruppen.

Tabelle 22: Verteilung der Mathematikleistungen bei Post-Test, Follow-up 1 und Follow-up 2 für die Gesamtstichprobe und die drei Stichprobengruppen unter Berücksichtigung aller vorhandenen Daten

	Gesamtstichprobe	**$\text{Gruppe}^{\text{Begl}}$**	**$\text{Gruppe}^{\text{Mat}}$**	**$\text{Gruppe}^{\text{Kontr}}$**
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Mathematikleistung Post-Test	25.49 (9.03)	24.15 (9.36)	25.75 (9.34)	26.45 (8.31)
Mathematikleistung Follow-up 1	27.62 (8.74)	26.65 (9.15)	28.04 (8.72)	28.12 (8.32)
Mathematikleistung Follow-up 2	44.01 (10.50)	42.78 (10.91)	44.14 (10.69)	44.98 (9.83)

Abschliessend soll in diesem Kapitel noch untersucht werden, ob sich die 45 Schülerinnen und Schüler, für die aufgrund des Alters von mehr als 113 Monaten keine genauen IQ-Werte bestimmt, sondern nur auf Basis der höchsten vorhandenen Altersnorm geschätzt werden konnten, von den Kindern mit regulär berechneten IQ-Werten in wichtigen Merkmalen unterscheiden. Damit soll geklärt werden, ob ein Weglassen dieser Kinder evtl. zu verzerrten Ergebnissen führen könnte. Wie der Tabelle 23 entnommen werden kann, unterscheiden sich die Mittelwerte und die Häufigkeiten bei allen aufgelisteten Merkmalen in den beiden Gruppen sehr stark.

Tabelle 23: Verteilung zentraler Merkmale bei Schülerinnen und Schülern (SuS) mit regulärer IQ-Bestimmung und bei SuS mit geschätzten IQ-Werten

	SuS mit regulärer IQ-Bestimmung	**SuS mit geschätzten IQ-Werten, auf Basis der höchsten Altersnorm**
	M (SD)	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.36 (6.19)	18.29 (7.11)
Alter (in Monaten)	103.20 (4.57)	117.73 (2.78)
IQ	105.62 (14.03)	96.13 (11.29)
SES	3.89 (1.14)	3.04 (1.31)
	n (%)	*n* (%)
Mädchen	423 (51.2)	14 (31.1)
Knaben	403 (48.8)	31 (68.9)
Förderbedarf Mathematik	158 (19.1)	18 (40.0)
Kinder mit Einschränkungen im Deutschverstehen	184 (22.3)	23 (51.1)

Es handelt sich bei den Schülerinnen und Schülern mit IQ-Werten, die auf Basis der höchsten Altersnorm geschätzt wurden, somit um Kinder, die über weniger mathematisches Vorwissen verfügen, deren IQ niedriger ausfällt und bei denen ein höherer Anteil an Kindern einen Förderbedarf in Mathematik oder Einschränkungen im Deutschverstehen aufweist. Da die Unterschiede rein deskriptiv bereits als sehr hoch bewertet werden können, wird an dieser Stelle auf eine statistische Prüfung der Unterschiede verzichtet und direkt angenommen, dass ein Weglassen dieser Kinder zu verzerrten Ergebnissen führen könnte. Daher werden im Ergebniskapitel jeweils die Ergebnisse für die im Folgenden als Stichprobe^{alle} bezeichnete Stichprobe berichtet, in denen die Daten der 45 Schülerinnen und Schüler für die Berechnungen eingeschlossen werden, deren IQ-Werte ausgehend von der höchsten zur Verfügung stehenden Altersnorm geschätzt wurden. Zusätzlich werden im Ergebniskapitel auch die Ergebnisse für die im Folgenden als Stichprobe^{IQstreng} bezeichnete Stichprobe dargestellt, in denen die Daten der genannten 45 Kinder nicht berücksichtigt sind und in der daher von einer zuverlässigeren IQ-Schätzung ausgegangen werden kann.

In den nächsten beiden Kapiteln werden die Stichprobe^{alle} und die Stichprobe^{IQstreng}, die für die Hypothesenprüfung zum Einsatz kommen, näher beschrieben. Da die verwendete Statistiksoftware für die Berechnung von Mehrebenenanalysen (HLM; Raudenbush, Bryk & Congdon, 2013) bei fehlenden Werten auf Individualebene einen fallweisen Ausschluss vornimmt (Raudenbush, Bryk, Cheong, Congdon & Du Toit, 2011), basieren die Modelle, über die in den Ergebniskapiteln berichtet wird, je nach Messzeitpunkt auf einer unterschiedlichen Anzahl von

Schülerinnen und Schülern. Daher wird in den nächsten beiden Kapiteln zusätzlich auch aufgezeigt, wie die zentralen Merkmale in der Stichprobe^{alle} und in der Stichprobe^{IQstreng} bei den verschiedenen Messzeitpunkten (Post-Test, Follow-up 1 und Follow-up 2) verteilt sind. Auf eine erneute statistische Prüfung der Unterschiede zwischen den drei Stichprobengruppen wird hier jedoch verzichtet und auf die entsprechenden Analysen zur Gesamtstichprobe verwiesen.

5.5.2 Stichprobe^{alle}

Stichprobe^{alle}: Post-Test

Für die Stichprobe^{alle} sind beim Post-Test die vollständigen Daten von 851 Schülerinnen und Schülern vorhanden und können somit für die Mehrebenenanalysen berücksichtigt werden (vgl. Tabelle 24). Es zeigt sich, dass die Werte für die Stichprobe^{alle} sehr nahe an den Werten liegen, die in Kapitel 5.5.1 für die Gesamtstichprobe ohne fallweisen Ausschluss bei fehlenden Werten berichtet wurden. Es wird daher nur auf Abweichungen eingegangen, die weiteren Werte können direkt der Tabelle 24 entnommen werden. Beim Post-Test können in der Stichprobe^{alle} für die Gruppe^{Begl} etwas höhere Mittelwerte für das mathematische Vorwissen (M = 20.97, SD = 6.42) und niedrigere Mittelwerte für den IQ (M = 103.88, SD = 15.24) im Vergleich zur Gesamtstichprobe festgestellt werden. Niedrigere Werte finden sich für die Gruppe^{Kontr} bei der Mathematikleistung (M = 26.34, SD = 8.33) sowie leicht höhere Werte beim Förderbedarf Mathematik (17.2%) und beim Anteil Kinder mit Einschränkungen im Deutschverstehen (21.4%).

Tabelle 24: Verteilung erfasster Merkmale für die Stichprobe^{alle} beim Post-Test über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen (N = 851)	**Gruppe^{Begl}** (n = 266)	**Gruppe^{Mat}** (n = 277)	**Gruppe^{Kontr}** (n = 308)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.15 (6.29)	20.97 (6.42)	20.50 (6.45)	21.90 (5.96)
Mathematikleistung (Post-Test)	25.49 (8.98)	24.19 (9.29)	25.79 (9.27)	26.34 (8.33)
Alter (in Monaten)	103.96 (5.55)	103.76 (5.97)	104.86 (5.47)	103.32 (5.15)
IQ „alle“	105.07 (13.92)	103.88 (15.24)	104.28 (13.44)	106.81 (13.00)
SES	3.84 (1.16)	3.75 (1.22)	3.74 (1.23)	4.00 (1.03)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	425 (49.9)	139 (52.3)	123 (44.4)	163 (52.9)
Knaben	426 (50.1)	127 (47.7)	154 (55.6)	145 (47.1)
Förderbedarf Mathematik	175 (20.6)	48 (18.0)	74 (26.7)	53 (17.2)
Kinder mit Einschränkungen im Deutschverstehen	205 (24.1)	77 (28.9)	62 (22.4)	66 (21.4)

Stichprobe^{alle}: Follow-up 1

In der Stichprobe^{alle} sind beim Follow-up 1 für 856 Schülerinnen und Schüler die kompletten Daten vorhanden und können somit für die Mehrebenenmodelle beim Follow-up 1 berücksichtigt werden. Auch hier zeigt sich, dass ein grosser Teil der Werte (fast) identisch ist mit den Werten, die in Kapitel 5.5.1 für die Gesamtstichprobe aufgezeigt wurden (vgl. Tabelle 25). Rein deskriptiv sind die Werte für den IQ sowohl über alle drei Gruppen hinweg (M = 105.19, SD = 13.98) als auch in den drei Stichprobengruppen (Gruppe^{Begl}: M = 104.21, SD = 15.55; Gruppe^{Mat}: M = 104.15, SD = 13.26; Gruppe^{Kontr}: M = 106.96, SD = 12.98) beim Follow-up 1 in der Stichprobe^{alle} leicht höher als in der Gesamtstichprobe. Im Vergleich zu den in Kapitel 5.5.1 erwähnten Werten ergeben sich auch leicht höhere Mittelwerte in der Gruppe^{Begl} für das mathematische Vorwissen (M = 21.04, SD = 6.41) und die Mathematikleistung beim Follow-up 1 (M = 26.79, SD = 9.02) sowie ein minimal höherer Wert für den IQ (M = 104.21, SD = 15.55). Geringe Abweichungen im Sinne von leicht niedrigeren Werten finden sich zudem für den Förderbedarf in Mathematik in allen drei Stichprobengruppen (Gruppe^{Begl}: 17.7%; Gruppe^{Mat}: 26.6%; Gruppe^{Kontr}: 17.0%), nicht aber für den Mittelwert über alle drei Gruppen hinweg (20.3%).

Tabelle 25: Verteilung erfasster Merkmale für die Stichprobe^{alle} beim Follow-up 1 über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen (N = 856)	**Gruppe^{Begl}** (n = 271)	**Gruppe^{Mat}** (n = 274)	**Gruppe^{Kontr}** (n = 311)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.20 (6.28)	21.04 (6.41)	20.56 (6.45)	21.90 (5.96)
Mathematikleistung (Follow-up 1)	27.66 (8.64)	26.79 (9.02)	28.04 (8.64)	28.06 (8.28)
Alter (in Monaten)	103.94 (5.54)	103.80 (5.95)	104.85 (5.45)	103.27 (5.14)
IQ „alle“	105.19 (13.98)	104.21 (15.55)	104.15 (13.26)	106.96 (12.98)
SES	3.84 (1.16)	3.75 (1.22)	3.76 (1.22)	4.00 (1.03)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	430 (50.2)	143 (52.8)	123 (44.9)	164 (52.7)
Knaben	426 (49.8)	128 (47.2)	151 (55.1)	147 (47.3)
Förderbedarf Mathematik	174 (20.3)	48 (17.7)	73 (26.6)	53 (17.0)
Kinder mit Einschränkungen im Deutschverstehen	204 (23.8)	77 (28.4)	61 (22.3)	66 (21.2)

Stichprobe^{alle}: Follow-up 2

Für die Stichprobe^{alle} liegen beim Follow-up 2 noch vollständige Daten von 818 Schülerinnen und Schülern vor, die für die Mehrebenenanalysen berücksichtigt werden können (vgl. Tabelle 26). Ein deskriptiver Vergleich zeigt, dass das mathematische Vorwissen in der Stichprobe^{alle} im Vergleich zu den Werten für die Gesamtstichprobe (vgl. Kap. 5.5.1) sowohl über alle drei Gruppen hinweg (M = 21.34, SD = 6.26) als auch in den drei Stichprobengruppen (Gruppe^{Begl}: M = 21.09, SD = 6.45; Gruppe^{Mat}: M = 20.61, SD = 6.42; Gruppe^{Kontr}: M = 22.21, SD = 5.87) leicht höher ausfällt. Auch bezüglich der Mathematikleistung beim Follow-up 2 gibt es deskriptiv leichte Abweichungen nach oben (Gruppe^{Mat}: M = 44.35, SD = 10.44) und nach unten (Gruppe^{Kontr}: M = 44.89, SD = 9.86). Beim IQ sind die Werte in allen Gruppen leicht höher (alle drei Gruppen: M = 105.32, SD = 14.06; Gruppe^{Begl}: M = 104.27, SD = 15.65; Gruppe^{Mat}: M = 104.29, SD = 13.42; Gruppe^{Kontr}: M = 107.17, SD = 12.94). Zusätzlich ist der Anteil Mädchen in der Gruppe^{Begl} etwas höher (53.3%) und in der Gruppe^{Mat} etwas niedriger (44.1%) als in der Gesamtstichprobe. Auch beim Förderbedarf in Mathematik gibt es einzelne Werte, die niedriger ausfallen (alle drei Gruppen: 19.8%; Gruppe^{Begl}: 17.2%; Gruppe^{Kontr}: 16.2%). Beim Anteil Kinder mit Einschränkungen im Deutschverstehen weist die Gruppe^{Mat} zudem leicht höhere Werte (23.0%) und die Gruppe^{Kontr} leicht niedrigere Werte (20.6%) auf.

Tabelle 26: Verteilung erfasster Merkmale für die Stichprobe^{alle} beim Follow-up 2 über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen (N = 818)	**Gruppe^{Begl}** (n = 261)	**Gruppe^{Mat}** (n = 261)	**Gruppe^{Kontr}** (n = 296)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.34 (6.26)	21.09 (6.45)	20.61 (6.42)	22.21 (5.87)
Mathematikleistung (Follow-up 2)	44.07 (10.37)	42.85 (10.80)	44.35 (10.44)	44.89 (9.86)
Alter (in Monaten)	103.91 (5.44)	103.90 (5.86)	104.75 (5.25)	103.17 (5.12)
IQ „alle“	105.32 (14.06)	104.27 (15.65)	104.29 (13.42)	107.17 (12.94)
SES	3.83 (1.16)	3.72 (1.22)	3.72 (1.22)	4.03 (1.01)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	409 (50.0)	139 (53.3)	115 (44.1)	155 (52.4)
Knaben	409 (50.0)	122 (46.7)	146 (55.9)	141 (47.6)
Förderbedarf Mathematik	162 (19.8)	45 (17.2)	69 (26.4)	48 (16.2)
Kinder mit Einschränkungen im Deutschverstehen	196 (24.0)	75 (28.7)	60 (23.0)	61 (20.6)

5.5.3 Stichprobe^{IQstreng}

Stichprobe^{IQstreng}: Post-Test

Beim Post-Test liegen für die Stichprobe^{IQstreng} für 806 Schülerinnen und Schüler vollständige Daten vor (vgl. Tabelle 27). Insgesamt sind die Abweichungen für die Stichprobe^{IQstreng} im Vergleich zur Gesamtstichprobe (vgl. Kap. 5.5.1) als grösser zu bewerten als bei der Stichprobe^{alle}. Beim mathematischen Vorwissen ergeben sich für die Stichprobe^{IQstreng} beim Post-Test teilweise leicht höhere Werte (alle drei Gruppen: $M = 21.31$, $SD = 6.21$; Gruppe^{Begl}: $M = 21.02$, $SD = 6.40$; Gruppe^{Mat}: $M = 20.89$, $SD = 6.30$) und die Mittelwerte für die Mathematikleistung sind beim Post-Test ebenfalls teilweise höher als in der Gesamtstichprobe (alle drei Gruppen: $M = 25.76$, $SD = 8.87$; Gruppe^{Begl}: $M = 24.37$, $SD = 9.10$; Gruppe^{Mat}: $M = 26.34$, $SD = 9.13$). Für das Alter sind überall niedrigere Werte festzustellen (alle drei Gruppen: $M = 103.19$, $SD = 4.58$; Gruppe^{Begl}: $M = 102.78$, $SD = 4.74$; Gruppe^{Mat}: $M = 103.98$, $SD = 4.40$; Gruppe^{Kontr}: $M = 102.85$, $SD = 4.52$). Beim IQ und beim SES ergeben sich hingegen nur minimale Unterschiede. Der Anteil Mädchen ist zudem in der Stichprobe^{IQstreng} über alle drei Gruppen hinweg (51.0%), in der Gruppe^{Begl} (53.8%) und insbesondere in der Gruppe^{Kontr} (54.4%) höher als in der Gesamtstichprobe. Der Anteil Kinder mit Einschränkungen im Deutschverstehen ist rein deskriptiv zudem für alle Gruppen niedriger (alle drei Gruppen: 22.6%; Gruppe^{Begl}: 27.7%; Gruppe^{Mat}: 19.7%; Gruppe^{Kontr}: 20.8%), wobei der grösste Unterschied auf die Gruppe^{Mat} entfällt.

Tabelle 27: Verteilung erfasster Merkmale für die Stichprobe^{IQstreng} beim Post-Test über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen ($N = 806$)	**Gruppe^{Begl}** ($n = 249$)	**Gruppe^{Mat}** ($n = 259$)	**Gruppe^{Kontr}** ($n = 298$)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.31 (6.21)	21.02 (6.40)	20.89 (6.30)	21.92 (5.93)
Mathematikleistung (Post-Test)	25.76 (8.87)	24.37 (9.10)	26.34 (9.13)	26.41 (8.33)
Alter (in Monaten)	103.19 (4.58)	102.78 (4.74)	103.98 (4.40)	102.85 (4.52)
IQ „streng“	105.57 (13.89)	104.39 (15.34)	104.85 (13.27)	107.19 (13.01)
SES	3.88 (1.14)	3.83 (1.17)	3.79 (1.21)	4.01 (1.03)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	411 (51.0)	134 (53.8)	115 (44.4)	162 (54.4)
Knaben	395 (49.0)	115 (46.2)	144 (55.6)	136 (45.6)
Förderbedarf Mathematik	157 (19.5)	43 (17.3)	63 (24.3)	51 (17.1)
Kinder mit Einschränkungen im Deutschverstehen	182 (22.6)	69 (27.7)	51 (19.7)	62 (20.8)

Stichprobe^{IQstreng}: Follow-up 1

Für den Follow-up 1 liegen für die Stichprobe^{IQstreng} die kompletten Daten von 811 Schülerinnen und Schülern vor, die für die Mehrebenenanalysen berücksichtigt werden können (vgl. Tabelle 28). Es zeigt sich, dass die Stichprobe^{IQstreng} beim Post-Test und beim Follow-up 1 über sehr ähnliche Kennwerte verfügt. Im Vergleich zu den in Kapitel 5.5.1 berichteten Werten für die Gesamtstichprobe ergeben sich somit sehr ähnliche Abweichungen, wie sie bereits beim Post-Test für die Stichprobe^{alle} berichtet wurden. Die einzelnen Werte über alle drei Gruppen hinweg sowie für die drei Stichprobengruppen können daher der Tabelle 28 entnommen werden, werden an dieser Stelle aber nicht weiter ausgeführt.

Tabelle 28: Verteilung erfasster Merkmale für die Stichprobe^{IQstreng} beim Follow-up 1 über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen ($N = 811$)	**Gruppe^{Begl}** ($n = 254$)	**Gruppe^{Mat}** ($n = 256$)	**Gruppe^{Kontr}** ($n = 301$)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.36 (6.19)	21.10 (6.38)	20.96 (6.29)	21.92 (5.93)
Mathematikleistung (Follow-up 1)	27.91 (8.52)	27.03 (8.83)	28.50 (8.45)	28.16 (8.27)
Alter (in Monaten)	103.18 (4.56)	102.84 (4.74)	103.95 (4.35)	102.80 (4.51)
IQ „streng“	105.69 (13.94)	104.72 (15.66)	104.71 (13.08)	107.34 (12.98)
SES	3.89 (1.13)	3.82 (1.18)	3.80 (1.20)	4.02 (1.03)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	416 (51.3)	138 (54.3)	115 (44.9)	163 (54.2)
Knaben	395 (48.7)	116 (45.7)	141 (55.1)	138 (45.8)
Förderbedarf Mathematik	156 (19.2)	43 (16.9)	62 (24.2)	51 (16.9)
Kinder mit Einschränkungen im Deutschverstehen	181 (22.3)	69 (27.2)	50 (19.5)	62 (20.6)

Stichprobe^{IQstreng}: Follow-up 2

Für die Stichprobe^{IQstreng} liegen noch die kompletten Daten von 778 Kindern vor, die für die Mehrebenenanalysen beim Follow-up 2 berücksichtigt werden können. Wie der Tabelle 29 entnommen werden kann, liegen die Werte in der Stichprobe^{IQstreng} für den Follow-up 2 sowohl über alle drei Gruppen hinweg als auch in den drei Stichprobengruppen in Bezug auf das mathematische Vorwissen, die Mathematikleistung beim Follow-up 2, den IQ und den SES etwas höher im Vergleich zu den im Kapitel 5.5.1 berichteten Werten für die Gesamtstichprobe, das

Alter hingegen ist etwas niedriger. Auch beim Geschlechteranteil gibt es leichte Veränderungen, sodass in der Stichprobe^{IQstreng} teilweise höhere Werte für den Anteil Mädchen resultieren (alle drei Gruppen: 50.9%; Gruppe^{Begl}: 54.7%; Gruppe^{Kontr}: 53.7%), für die Gruppe^{Mat} ist dieser Wert hingegen niedriger (43.9%) als in der Gesamtstichprobe. Auch der Anteil Kinder mit Förderbedarf und der Anteil Kinder mit Einschränkungen im Deutschverstehen weichen von den Werten für die Gesamtstichprobe ab. So ist der Anteil Kinder mit Förderbedarf in allen Gruppen niedriger (alle drei Gruppen: 18.9%; Gruppe^{Begl}: 16.7%; Gruppe^{Mat}: 24.4%; Gruppe^{Kontr}: 16.0%), ebenso wie der Anteil Kinder mit Einschränkungen im Deutschverstehen (alle drei Gruppen: 22.4%; Gruppe^{Begl}: 27.3%; Gruppe^{Mat}: 20.3%; Gruppe^{Kontr}: 19.9%).

Tabelle 29: Verteilung erfasster Merkmale für die Stichprobe^{IQstreng} beim Follow-up 2 über alle drei Gruppen hinweg und separat für die drei Stichprobengruppen

	Alle drei Gruppen (N = 778)	**Gruppe^{Begl}** (n = 245)	**Gruppe^{Mat}** (n = 246)	**Gruppe^{Kontr}** (n = 287)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Vorwissen Mathematik (Pre-Test)	21.48 (6.19)	21.14 (6.42)	20.92 (6.31)	22.25 (5.82)
Mathematikleistung (Follow-up 2)	44.34 (10.29)	42.97 (10.74)	44.80 (10.33)	45.11 (9.77)
Alter (in Monaten)	103.20 (4.54)	102.98 (4.73)	103.98 (4.33)	102.72 (4.49)
IQ „streng“	105.83 (14.03)	104.79 (15.73)	104.86 (13.31)	107.55 (12.93)
SES	3.88 (1.13)	3.80 (1.18)	3.77 (1.20)	4.05 (1.01)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Mädchen	396 (50.9)	134 (54.7)	108 (43.9)	154 (53.7)
Knaben	382 (49.1)	111 (45.3)	138 (56.1)	133 (46.3)
Förderbedarf Mathematik	147 (18.9)	41 (16.7)	60 (24.4)	46 (16.0)
Kinder mit Einschränkungen im Deutschverstehen	174 (22.4)	67 (27.3)	50 (20.3)	57 (19.9)

Im Ergebniskapitel werden jeweils zunächst Modelle für die Stichprobe^{alle} präsentiert, in denen Schülerinnen und Schüler berücksichtigt werden, deren IQ auf Basis der höchsten zur Verfügung stehenden Altersnorm geschätzt wurde. Es werden somit in der Stichprobe^{alle} auch die Daten von 45 Schülerinnen und Schüler verwendet, die im Durchschnitt über weniger mathematisches Vorwissen verfügen, deren IQ niedriger ausfällt und bei denen ein höherer Anteil an Kindern einen Förderbedarf in Mathematik oder Einschränkungen im Deutschverstehen aufweist (vgl. Tabelle 23). Zusätzlich wird auch berichtet, wie die entsprechenden Modelle für die Stichprobe^{IQstreng} ausfallen, in der die Daten der genannten 45 Kinder nicht berücksichtigt werden. Die Ergebnisse für die Stichprobe^{alle} sind somit repräsentativer mit Blick auf das untere Leistungsspektrum, in der Stichprobe^{IQstreng} können hingegen zuverlässigere IQ-Schätzungen angenommen werden.

5.6 Auswertungsmethoden

Im Folgenden werden zunächst die statistischen Auswertungsmethoden für die Beschreibung der Stichproben dargestellt, die in den vorangehenden Kapiteln präsentiert wurde, und anschliessend wird ausführlich auf die verwendeten Mehrebenenanalysen eingegangen, die für die Beantwortung der Fragestellungen der vorliegenden Untersuchung eingesetzt werden.

5.6.1 Statistische Analysen für die Beschreibung der Stichprobe

Alle statistischen Analysen in dieser Arbeit wurden mit Ausnahme der Mehrebenenanalysen mit SPSS (Version 23.0.0.2) erstellt. Für die Analyse von Unterschieden zwischen den Stichprobengruppen in relevanten, intervallskalierten Merkmalen (Vorwissen Mathematik, IQ, SES und Alter) wurden einfaktorielle Varianzanalysen berechnet (vgl. Kap. 5.5). Diese eignen sich, um die Mittelwerte von mehr als zwei unabhängigen Stichprobengruppen miteinander zu vergleichen (Field, 2009). Da einfaktorielle Varianzanalysen nur für die Analyse von Unterschieden zwischen den Stichprobengruppen im Rahmen der Stichprobenbeschreibung verwendet werden, wird diese Auswertungsmethode in dieser Arbeit nicht ausgeführt, sondern auf die entsprechende Fachliteratur verwiesen (z.B. Akremi, Baur & Fromm, 2011; Field, 2009; Tabachnick & Fidell, 2007). Ausführlicher diskutiert werden im Folgenden jedoch die Prüfung der Voraussetzungen für die vorliegende Untersuchung und die daraus zu ziehenden Folgerungen für die zu wählenden Testverfahren.

Zentrale Voraussetzungen für die einfaktorielle Varianzanalyse sind normalverteilte Variablen und Varianzhomogenität (Field, 2009). Diese Voraussetzungen sind bei den Variablen Vorwissen Mathematik, IQ, SES und Alter nicht immer gegeben und werden daher zunächst für die Normalverteilung und im Anschluss für die Varianzhomogenität diskutiert. Gemäss Lienert und Raatz (1998, S. 148) kann davon ausgegangen werden, dass bei Werten für die Schiefe kleiner oder gleich 0.5 und Kurtosis-Werten kleiner oder gleich 1.0 trotz Verletzung der Normalverteilungsannahme keine nennenswerten Verzerrungen der Ergebnisse vorliegen sollten. Gemäss Curran, West und Finch (1996, S. 26) kann aufgrund von Monte-Carlo-Computersimulationen sogar angenommen werden, dass dies auch auf Werte für die Schiefe von kleiner oder gleich 2.0 und auf Werte der Kurtosis von kleiner oder gleich 4.0 zutrifft. Diese Faustregeln sind wichtig, da die klassischen Testverfahren für die Prüfung der Normalverteilung (Kolmogorov-Smirnov-Test oder Shapiro-Wilk-Test) bei grossen Stichproben auch bei minimalen Abweichungen von der Normalverteilung signifikante Ergebnisse zeigen (Field, 2009). Für die vorliegende Untersuchung kann trotz gewisser Abweichungen von der Normalverteilung somit für das mathematische Vorwissen (Schiefe = –0.058, Kurtosis = –0.782) und für den IQ (Schiefe = 0.223, Kurtosis = –0.088) aufgrund der strengeren Faustregeln von Lienert und Raatz (1998) angenommen werden, dass die Ergebnisse von einfaktoriellen Varianzanalysen zuverlässig sind. Für den SES (Schiefe = –0.439, Kurtosis = –1.077) und das Alter (Schiefe = 0.579, Kurtosis = 0.119) kann auf Basis der etwas weiteren Faustregeln von Curran et al. (1996) die gleiche Annahme getroffen werden.

Bei Verletzung der Varianzhomogenität kann für die Feststellung von Gruppenunterschieden bei der einfaktoriellen Varianzanalyse der Welch-Test verwendet werden, der robust ist für entsprechende Verletzungen (Field, 2009). In der hier vorliegenden Untersuchung ist dies beim IQ für beide IQ-Variablen (IQ „alle“: Levene-Test: $F[2, 868] = 3.215$, $p = .041$; IQ „streng“: Levene-Test: $F[2, 823] = 3.424$, $p = .033$) und beim SES (Levene-Test: $F[2,884] = 13.447$, $p < .000$) der Fall. Aufgrund der festgestellten Varianzheterogenität wird für die Post-hoc-Vergleiche beim IQ und beim SES zudem der Games-Howell-Test verwendet, da dieser trotz Verletzung der Varianzhomogenität zu zuverlässigen Ergebnissen führt (Field, 2009). Für die restlichen Variablen wird für die Post-hoc-Vergleiche der Tukey-Kramer-Test für ungleiche Stichprobengrössen verwendet (Field, 2009).

Bei den kategorialen Variablen (Geschlecht, Förderbedarf in Mathematik und Einschränkungen im Deutschverstehen) wurden für die Feststellung allfälliger Gruppenunterschiede χ^2-Tests nach Pearson (Kreuztabellen) berechnet (vgl. Kap. 5.5). Für keine der genannten Variablen ist die Voraussetzung verletzt, wonach in jeder Zelle mindestens fünf Fälle vorhanden sein müssen (Field, 2009).

5.6.2 Mehrebenenanalysen mit HLM

Für die Beantwortung der Fragestellungen der vorliegenden Untersuchung werden Mehrebenenanalysen verwendet. Daher wird im Folgenden ausführlich auf diese Auswertungsmethode eingegangen, und es wird thematisiert, was bei der Anwendung dieser Methode beachtet werden muss.

Gründe für die Verwendung von Mehrebenenanalysen

Mehrebenenanalysen (auch als Mehrebenenmodelle, hierarchische lineare Modelle, random coefficient models und mixed models bekannt) sind Analyseverfahren, die sich für die Auswertung von hierarchisch strukturierten Daten eignen (Geiser, 2011). Mit Bezug auf den Schulkontext bedeutet dies, dass Daten der Schülerinnen und Schüler „[…] in Schulklassen

‚geschachtelt‘ […] sind“ (Geiser, 2011, S. 199). Somit kann bei solchen Daten die für andere statistische Analyseverfahren wichtige Voraussetzung der Unabhängigkeit der Beobachtungen nicht erfüllt werden, was zu gravierenden Verzerrungen der Ergebnisse führen kann (Geiser, 2011). Die Abhängigkeit der verschiedenen Beobachtungen voneinander zeigt sich dadurch...

> [...] dass [Schülerinnen und] Schüler innerhalb ein und derselben Schulklasse einander in Bezug auf bestimmte Merkmale (z.B. Schulleistung) ähnlicher sind als [Schülerinnen und] Schüler, die zu einer anderen Klasse gehören, da [Schülerinnen und] Schüler derselben Klasse beispielsweise bestimmten gemeinsamen Einflüssen ausgesetzt sind (gemeinsamer Unterricht, soziale Interaktionen innerhalb der Klasse etc.). (Geiser, 2011, S. 199)

Diese Datenstruktur kann in Mehrebenenmodellen spezifisch modelliert werden. Dabei werden im Schulkontext auf Individualebene (auch Level 1 oder Mikroebene genannt) die Daten der Schülerinnen und Schüler und auf Klassenebene (auch Level 2 oder Makroebene genannt) klassenspezifische Merkmale berücksichtigt, wobei sich grundsätzlich auch noch weitere Ebenen hinzufügen lassen (z.B. Schulhaus oder Schulgemeinde; Geiser, 2011). Mehrebenenanalysen berücksichtigen dazu sowohl die Variabilität, die zwischen Individuen, als auch die Variabilität, die zwischen Gruppen bzw. Klassen besteht (Snijders & Bosker, 2012). Die Vorteile der Mehrebenenanalysen lassen sich folgendermassen zusammenfassen:

- Adäquate Berücksichtigung der durch die hierarchische Datenstruktur (Clusterung) entstehenden Abhängigkeiten in den Daten
- Vermeidung von
    - Standardfehlerbias (zu gering geschätzte Standardfehler)
    - inflationiertem Alpha-Fehler-Risiko (zu liberale Signifikanztests)
    - zu eng geschätzten Konfidenzintervallen
- Betrachtung von Einflussfaktoren auf Mikro- und Makroebene sowie Möglichkeit zur Analyse von Cross-Level-Interaktionen
- Hohe Flexibilität und weniger strenge Voraussetzungen als bei anderen statistischen Verfahren (z.B. Varianzanalyse mit Messwiederholungen). (Geiser, 2011, S. 200)

Mehrebenenanalysen bauen auf linearen Regressionsanalysen auf, wobei die Regressionskonstante (sowie in Modellen mit Random Slopes auch die Slopes) „[…] über die Kontexteinheiten hinweg variieren darf“ (Langer, 2009, S. 109). Es gibt verschiedene Softwareprogramme, die Mehrebenenanalysen berechnen können. Für die vorliegende Arbeit wurde die Statistiksoftware HLM von Raudenbush et al. (2013) in der Version 7.01 verwendet.

Regressionsgleichungen der Mehrebenenanalysen

Bei Mehrebenenanalysen wird jeweils als Erstes ein sogenanntes Nullmodell (auch unkonditioniertes Modell oder Intercept-only Model genannt) erstellt. Das Nullmodell ist ein Mehrebenenmodell, bei dem nur die abhängige Variable ohne Prädiktoren eingefügt wird. Die Gleichung für das Nullmodell auf Individualebene lautet gemäss Hox (2010, S. 15):[19]

$$Y_{ij} = \mathrm{\beta}_{0j} + e_{ij} \tag{1.1}$$

Y steht für den Wert der abhängigen Variable, j bezeichnet hier die Schulklasse ($j = 1 \ldots J$) und i den einzelnen Schüler bzw. die Schülerin ($i = 1 \ldots \mathrm{n_j}$). $\mathrm{\beta}_{0j}$ steht für den Intercept (Achsenabschnitt) der Regressionsgleichung, und e_{ij} bezeichnet den Fehlerterm auf Individualebene (auch Residualwert genannt) der Gleichung. Die Gleichung für das Nullmodell auf Klassenebene lautet gemäss Hox (2010, S. 15):

$$\mathrm{\beta}_{0j} = \mathrm{\gamma}_{00} + u_{0j} \tag{1.2}$$

$\mathrm{\gamma}_{00}$ steht in dieser Gleichung für den Regressionskoeffizienten des Intercepts und u_{0j} für den Fehlerterm auf Klassenebene (auch Residualwert genannt). Die beiden Gleichungen 1.1 und 1.2 können auch in einer Gleichung gemeinsam notiert werden (Hox, 2010, S. 15):

$$Y_{ij} = \mathrm{\gamma}_{00} + u_{0j} + e_{ij} \tag{1.3}$$

Mithilfe des Nullmodells kann die sogenannte Intraklassen-Korrelation ($\mathrm{ICC_1}$) berechnet werden, die angibt, welcher Anteil der Varianz des Modells durch Gruppenunterschiede erklärt werden kann (Hox, 2010, S. 15):

$$\mathrm{\rho} = \mathrm{\sigma}^2_{u0} / \mathrm{\sigma}^2_{u0} + \mathrm{\sigma}^2_{e} \tag{1.4}$$

Mit $\mathrm{\sigma}^2_{u0}$ wird die Varianz des Fehlerterms auf Klassenebene (u_{0j}) bezeichnet, $\mathrm{\sigma}^2_{e}$ steht für die Varianz des Fehlerterms auf Individualebene (e_{ij}). Bereits bei einem $\mathrm{ICC_1}$ ab einem Wert von 0.10 (Muthén, 1997, 457f) kann die α-Fehlerquote ohne Berücksichtigung der hierarchischen Datenstruktur stark ansteigen, sodass Mehrebenenanalysen angezeigt sind.

Aufbauend auf dem Nullmodell, werden in der vorliegenden Untersuchung gemäss Empfehlungen von Hox (2010) schrittweise und theoriegeleitet Prädiktoren zuerst auf Individual- (X_{ij}) und dann auf Klassenebene (Z_j) eingefügt, um deren Einfluss auf die abhängige Variable zu

[19] Änderungen im Vergleich zum Original: Im Buch von Hox (2010) werden bei allen Gleichungen, die im Folgenden zitiert werden, die griechischen Buchstaben kursiv geschrieben. Für die vorliegende Arbeit wurde auf eine Kursivsetzung griechischer Buchstaben verzichtet (vgl. Deutsche Gesellschaft für Psychologie, 2016).

prüfen. Gemäss den Ausführungen im Kapitel 2.3 müssen aus theoretischer Perspektive als Einflussfaktoren für die Mathematikleistung insbesondere das Vorwissen, der IQ, das Geschlecht, ein Förderbedarf in Mathematik sowie Einschränkungen im Deutschverstehen auf Individual- und Klassenebene berücksichtigt werden.[20] In den so berechneten Modellen werden zunächst nur die Intercepts über die Schulklassen hinweg als zufällige Effekte definiert. Die mathematische Gleichung auf Individualebene eines solchen als Random Intercept bezeichneten, Modells soll hier am Beispiel des Modells 1 für den Post-Test der Stichprobe[alle] (vgl. Tabelle 37) dargestellt werden (vgl. Hox, 2010):

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \beta_{5j}X_{5ij} + e_{ij} \qquad (1.5)$$

Unter Verwendung der Variablennamen lautet die obige Gleichung:

$$\text{Mathematikleistung (Post-Test)} = \beta_{0j} + \beta_{1j}\text{Vorwissen Mathematik}_{ij} + \beta_{2j}\text{IQ}_{ij} + \beta_{3j}\text{Geschlecht}_{ij} + \beta_{4j}\text{Förderbedarf Mathematik}_{ij} + \beta_{5j}\text{Einschränkungen Deutschverstehen}_{ij} + e_{ij}$$

Für die Klassenebene lassen sich folgende Gleichungen für das Modell 1 aus der vorliegenden Arbeit formulieren, wobei der Term u_{0j} darauf verweist, dass hier nur der Intercept β_{0j} zwischen den Klassen variiert (vgl. Tabelle 37):

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} + \gamma_{01}Z_j + u_{0j} \\
\beta_{1j} &= \gamma_{10} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30} \\
\beta_{4j} &= \gamma_{40} \\
\beta_{5j} &= \gamma_{50} \qquad (1.6)
\end{aligned}$$

Werden wiederum die Variablennamen verwendet, lautet auf Klassenebene die Gleichung, in der auf Klassenebene der Prädiktor „IQ-Klassenmittelwert" eingefügt wurde, für β_{0j} wie folgt (die restlichen Gleichungen auf Level 2 werden hier nicht nochmals dargestellt):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{IQ-Klassenmittelwert}_j + u_{0j} \qquad (1.7)$$

Auch für das Random-Intercept-Modell lässt sich eine gemeinsame Gleichung für beide Ebenen formulieren (hier nur mit Variablennamen dargestellt):

[20] Im Rahmen der vorliegenden Untersuchung konnten aus organisatorischen Gründen keine kognitiven Faktoren (z.B. Arbeitsgedächtnis) und keine qualitativen Unterrichtsmerkmale erhoben werden. Diese Einflussfaktoren können daher auch nicht für die Analysen berücksichtigt werden.

Mathematikleistung (Post-Test) = γ_{00} + γ_{01}IQ-Klassenmittelwert$_j$ + γ_{10}Vorwissen Mathematik$_{ij}$ + γ_{20}IQ$_{ij}$ + γ_{30}Geschlecht$_{ij}$ + γ_{40}Förderbedarf Mathematik$_{ij}$ + γ_{50}Einschränkungen Deutschverstehen$_{ij}$ + u_{0j} + e_{ij} (1.8)

Nach Prüfung aller infrage kommenden Individual- und Klassenprädiktoren wird in einem nächsten Schritt untersucht, ob zusätzlich auch die Slopes über die Klassen hinweg variieren (Hox, 2010). Die Gleichung auf Individualebene bleibt dabei dieselbe, allerdings müssen die Gleichungen für die Klassenebene angepasst werden, was hier am Beispiel einer Random Slope für den Förderbedarf in Mathematik dargestellt wird. Die Gleichung für β_{4j} wird hier also um den Fehlerterm u_{4j} ergänzt, der entsprechend auch in die gemeinsame Gleichung für beide Ebenen aufgenommen wird (hier nicht dargestellt):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + u_{4j}$$

$$\beta_{5j} = \gamma_{50} \qquad (1.9)$$

In einem weiteren Schritt wird gemäss Hox (2010) geprüft, ob sich die Unterschiede in den Slopes zwischen den Klassen durch Klassenmerkmale erklären lassen. Ein solches Modell wird als Cross-Level-Modell bezeichnet. In der vorliegenden Arbeit wird im Modell 3 (vgl. Tabelle 37) eine Cross-Level-Interaktion zwischen Förderbedarf in Mathematik und Anteil Schülerinnen und Schüler mit Einschränkungen im Deutschverstehen in den Klassen präsentiert, sodass beispielhaft für dieses Modell folgende Gleichung für die Individualebene formuliert werden kann (vgl. Hox, 2010):

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + \beta_{5j}X_{5ij} + e_{ij} \qquad (1.10)$$

Auf Klassenebene lauten die Gleichungen (vgl. Hox, 2010):

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41}Z_j + u_{4j}$$

$\beta_{5j} = \gamma_{50}$ (1.11)

Die Gleichung für beide Ebenen lautet demnach (vgl. Hox, 2010):

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{20}X_{2ij} + \gamma_{30}X_{3ij} + \gamma_{40}X_{4ij} + \gamma_{41}X_{4ij}Z_j + \gamma_{50}X_{5ij} + u_{4j}X_{4ij} + u_{0j} + e_{ij} \quad (1.12)$$

Unter Verwendung der Variablennamen lautet die obige Gleichung:

$Y_{ij} = \gamma_{00} + \gamma_{10}$Vorwissen Mathematik$_{ij}$ + γ_{20}IQ$_{ij}$ + γ_{30}Geschlecht$_{ij}$ + γ_{40}Förderbedarf Mathematik$_{ij}$ + γ_{41}Förderbedarf Mathematik$_{ij}$ × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen$_j$ + γ_{50} Einschränkungen im Deutschverstehen$_{ij}$ + u_{4j}Förderbedarf Mathematik$_{ij}$ + $u_{0j} + e_{ij}$

Standardisierte vs. unstandardisierte Regressionskoeffizienten

Für die oben genannten Gleichungen wurden die Regressionskoeffizienten jeweils mit β bezeichnet, wie dies in der entsprechenden Fachliteratur üblich ist (Hox, 2010; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Damit sind jeweils standardisierte Regressionskoeffizienten gemeint. Das Statistikprogramm HLM (Raudenbush et al., 2013) hingegen berechnet unstandardisierte Regressionskoeffizienten, die mit *B* bezeichnet werden. Eine Berechnung von standardisierten Regressionskoeffizienten ist mit HLM nur möglich, wenn die Variablen im Voraus standardisiert werden (Raudenbush et al., 2011). Alternativ können jedoch nachträglich standardisierte Regressionskoeffizienten mit folgender Gleichung manuell berechnet werden (Hox, 2010, S. 22):

Standardisierter Regressionskoeffizient = (unstandardisierter Regressionskoeffizient des Prädiktors × Standardabweichung des Prädiktors) / Standardabweichung der abhängigen Variablen (1.13)

Bei unstandardisierten Regressionskoeffizienten wird die natürliche Metrik der Variablen beibehalten (Hox, 2010), was für die Interpretation der Ergebnisse mit Bezug auf die abhängigen Variablen (Leistung im Mathematiktest) in der vorliegenden Arbeit von Vorteil ist. Zudem können so die Regressionskoeffizienten für die zwei unterschiedlichen Untersuchungsstichproben (Stichprobe$^{\text{alle}}$ und Stichprobe$^{\text{IQstreng}}$; vgl. Kap. 5.5) miteinander verglichen werden (Hox, 2010). Daher werden in dieser Arbeit bei der Ergebnispräsentation im Kapitel 6 in erster Linie unstandardisierte Regressionskoeffizienten aufgezeigt. Da ein Vergleich der Regressionskoeffizienten zwischen verschiedenen Messzeitpunkten und zwischen verschiedenen Prädiktoren des gleichen Modells bzw. für die Diskussion auch ein Vergleich mit anderen Studien für gewisse Variablen dennoch interessant ist (insbesondere IQ und mathematisches Vorwissen), werden für einzelne intervallskalierte Variablen zusätzlich zu den unstandardisierten Regressionskoeffizienten auch nachträglich standardisierte Regressionskoeffizienten berechnet und im Rahmen der Ergebnisse präsentiert (vgl. Kap. 6). Bei nominalskalierten Variablen sowie bei Dummy-codierten Variablen können standardisierte Regressionskoeffizienten hingegen aufgrund der Metrik der Variablen nicht sinnvoll interpretiert werden, sodass für diese Variablen keine standardisierten Regressionskoeffizienten berichtet werden (vgl. Wolf & Best, 2010).

Berechnungsverfahren für die Regressionskoeffizienten und die Zentrierung von Prädiktoren

Es gibt verschiedene Methoden für die Berechnung der Regressionskoeffizienten in Mehrebenenanalysen. Die am häufigsten verwendete Methode ist die Maximum-Likelihood-Methode (ML), da diese sehr robust und weniger anfällig für milde Verletzungen von methodischen Voraussetzungen ist (Hox, 2010; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Dabei können zwei unterschiedliche Likelihood-Funktionen zur Anwendung kommen: Die Full Maximum Likelihood (FML) und die Restricted Maximum Likelihood (RML; Hox, 2010). In der vorliegenden Arbeit werden Regressionskoeffizienten berichtet, die mit dem RML-Verfahren berechnet werden, da dieses weniger anfällig für Verletzungen der Normalverteilung der Residualvarianzen ist (Greenberg & Philips, 2013; Hox, 2010). Das FML-Verfahren wird hingegen für die Berechnung von Devianzwerten verwendet, die für den Vergleich zweier Modelle benötigt werden (vgl. nächster Abschnitt „*Verfahren zur Überprüfung von Modellverbesserungen*“). Zusätzlich muss beim Berechnen von Mehrebenenmodellen beim Einfügen von Individualprädiktoren eine von drei gängigen Zentrierungsmethoden gewählt werden, was weitreichende Konsequenzen für die Interpretation der Regressionskoeffizienten hat (z.B. Enders & Tofighi, 2007; Nezlek, Schröder-Abé & Schütz, 2006). Für Dummy-Variablen wurde in der vorliegenden Arbeit auf Individualebene eine Zentrierung am Nullpunkt gewählt, da bei diesen Variablen der Nullpunkt sinnvoll interpretiert werden kann (z.B. Förderbedarf in Mathematik: Nein = 0, Ja = 1; Enders & Tofighi, 2007). In Modellen mit Prädiktoren auf Individual- und auf Klassenebene kommt für Individualprädiktoren, bei denen der Nullpunkt nicht sinnvoll zu interpretieren ist (z.B. Alter, IQ), grundsätzlich sowohl eine Zentrierung am Gesamtmittelwert (Grand-Mean-Zentrierung) als auch eine Zentrierung am Gruppenmittelwert (Group-Mean-Zentrierung) infrage (Enders & Tofighi, 2007). Bei der Zentrierung am Gesamtmittelwert wird jeweils der Gesamtmittelwert vom individuellen Wert subtrahiert, bei der Zentrierung am Gruppenmittelwert wird hingegen der Gruppenmittelwert (d.h. der Mittelwert der Klasse) vom individuellen Wert abgezogen (Enders & Tofighi, 2007; Hox, 2010; Snijders & Bosker, 2012). Gemäss Lüdtke, Robitzsch, Trautwein und Kunter (2009) sollten bei Variablen wie dem SES, bei dem die aggregierte Level-2-Variable als Folge der Level-1-Variable anzusehen ist (aggregierter SES ist abhängig vom individuellen SES), eine Grand-Mean-Zentrierung gewählt werden.[21] Da alle Prädiktorvariablen dieser Aggregationslogik entsprechen, wird in der vorliegenden Untersuchung für alle Individualprädiktoren mit Ausnahme der Dummy-Variablen eine Zentrierung am Gesamtmittelwert gewählt. Für Prädiktoren auf Klassenebene besteht nur die Wahl zwischen keiner Zentrierung und der Zentrierung am Gesamtmittelwert, wobei Letztere in dieser Untersuchung dann angewendet wird, wenn der Wert null einer Variablen nicht sinnvoll interpretiert werden kann (Nezlek et al., 2006).

Verfahren zur Überprüfung von Modellverbesserungen

Wie vorgängig geschildert, werden für die vorliegende Arbeit die berichteten Modelle schrittweise aufgebaut (Nullmodell → Random-Intercept-Modelle → Random-Slope-Modelle → Cross-Level-Modelle), wobei jeweils zunächst Individualprädiktoren und dann Klassenprädiktoren theoriegeleitet eingefügt werden (Hox, 2010). Um zu entscheiden, ob ein Modell mit einem zusätzlichen Prädiktor im Vergleich zum Modell ohne diesen Prädiktor als besser zu den realen Daten passend bezeichnet werden kann, steht in HLM (Raudenbush et al., 2013) ein Testverfahren zur Verfügung, das für die vorliegende Untersuchung genutzt wird. Dafür werden die Devianzwerte der zu vergleichenden Modelle benötigt, die HLM (Raudenbush et al., 2013) standardmässig berechnet. Je niedriger der Devianzwert eines Modells ist, umso besser passt es zu den Daten (Modell-Fit; Raudenbush & Bryk, 2002). Entsprechend wird im Testverfahren mithilfe einer χ^2-Verteilung geprüft, ob die Differenz der Devianzwerte (Devianz des Modells ohne Prädiktor – Devianz des Modells mit Prädiktor) signifikant ist und somit das Modell mit zusätzlichem Prädiktor besser zu den Daten passt (Raudenbush & Bryk, 2002). Wenn ja, wird in der vorliegenden Untersuchung der jeweilige Prädiktor im Modell belassen und das Hinzufügen weiterer Prädiktoren in gleicher Weise untersucht, bis keine Modellverbesserung mehr erreicht werden kann. Anschliessend wird das gleiche Verfahren zur Prüfung von Modellverbesserungen durch das Einfügen von Random Slopes eingesetzt usw.

Varianzaufklärung in Mehrebenenmodellen und die Berechnung von Effektstärken

Ein weiteres wichtiges Mass für die Beurteilung von Mehrebenenmodellen ist die Varianzaufklärung des jeweiligen Modells auf Individual- und auf Klassenebene (Hox, 2010; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Da die klassische Berechnungsweise der Varianzaufklärung, wie sie z.B. bei Raudenbush und Bryk (2002) und Hox (2010) zu finden ist, zu negativen Varianzaufklärungen führen kann, kann alternativ eine von Snijders und Bosker (1994) angepasste Formel für die Berechnung verwendet werden (vgl. Snijders & Bosker, 2012). Die-

[21] Eine Zentrierung am Gruppenmittelwert empfiehlt sich hingegen bei Prädiktoren auf Individualebene, die beeinflusst sind von Variablen auf Klassenebene (z.B. Motivation bei Hausaufgaben steht in Abhängigkeit zur Hausaufgabenqualität; Lüdtke et al., 2009).

se korrigierte Varianzaufklärung nach Snijders und Bosker (1994) wird auf Individualebene wie folgt berechnet:

$$R^2_{\text{Level 1}} = (\sigma^2_{e|\text{Nullmodell}} + \sigma^2_{u0|\text{Nullmodell}}) - (\sigma^2_{e|\text{aktuelles Modell}} + \sigma^2_{u0|\text{aktuelles Modell}}) \,/\, (\sigma^2_{e|\text{Nullmodell}} + \sigma^2_{u0|\text{Nullmodell}}) \tag{1.14}$$

Für die Klassenebene lautet die Gleichung (Snijders & Bosker, 1994):

$$R^2_{\text{Level 2}} = (\sigma^2_{e|\text{Nullmodell}} / n + \sigma^2_{u0|\text{Nullmodell}}) - (\sigma^2_{e|\text{aktuelles Modell}} / n + \sigma^2_{u0|\text{aktuelles Modell}}) \,/\, (\sigma^2_{e|\text{Nullmodell}} / n + \sigma^2_{u0|\text{Nullmodell}}) \tag{1.15}$$

Für ungleiche Gruppengrössen auf Klassenebene wird das harmonische Mittel zur Berechnung von n gewählt ($\{(1/N) \sum_{j} (1/n_j)\}^{-1}$; Snijders & Bosker, 1994, S. 354).

Die Berechnung der Varianzaufklärung von Modellen mit Random Slopes wird von Snijders und Bosker (1994) zwar ausgeführt, kann jedoch nicht nachvollzogen werden, da die Angaben unvollständig sind mit Bezug auf die zu verwendenden Schätzwerte im HLM-Output. Entsprechend wird in der vorliegenden Untersuchung bei Modellen mit Random Slopes der alternativen Empfehlung von Snijders und Bosker (1994) gefolgt. Demnach hat das Hinzufügen von Random Slopes kaum Auswirkungen auf die Varianzaufklärung, und es ist gerechtfertigt, die Varianzaufklärung auf Basis des Modells ohne Random Slopes zu berechnen (Snijders & Bosker, 1994, S. 358).

Zusätzlich zur Varianzaufklärung können auch Effektstärken für die Regressionskoeffizienten berechnet werden. Für Dummy-codierte Variablen lautet die Berechnung gemäss Tymms (2004, S. 57):

$$\Delta = \beta_1 / \sigma_e \tag{1.16}$$

Als Richtgrösse für die Einschätzung von Effektstärken wird meistens auf Cohen (1988) verwiesen, wonach Effektstärken von 0.20 als klein, solche von 0.50 als mittel und solche von 0.80 als hoch gross eingeschätzt werden können. Allerdings muss gerade bei Untersuchungen in natürlichen Kontexten wie der Schule beachtet werden, dass viele Faktoren die abhängige Variable beeinflussen, sodass die Erklärungskraft eines einzelnen Prädiktors bereits dadurch eingeschränkt wird (Fuchs, Fuchs & Compton, 2012; Trautwein, Niggli, Schnyder & Lüdtke, 2009).

> For this reason, small effect sizes are typically considered meaningful in this research field, especially if they are associated with teaching characteristics that are

> modifiable […]. We thus suggest that a small effect size of 0.20 should also be considered meaningful in the present research. (Trautwein et al., 2009, S. 181)

Voraussetzungen der Mehrebenenanalysen

Abschliessend wird im Folgenden die Prüfung von Voraussetzungen bei Mehrebenenanalysen thematisiert. Folgende Auflistung gibt eine Übersicht über zentrale zu prüfende Voraussetzungen bei Mehrebenenanalysen mit HLM (Raudenbush & Bryk, 2002):

a) Lineare Zusammenhänge zwischen Prädiktorvariablen und abhängiger Variable

b) Homogene Varianzen der Fehlerterme (σ^2_e) auf Individualebene

c) Normalverteilte Fehlerterme (e_{ij}) auf Individualebene

d) Normalverteilte Fehlerterme (u_{0j}) auf Klassenebene

In der Statistiksoftware HLM (Raudenbush et al., 2013) können sogenannte „residual files“ für die Individual- und die Klassenebene angefordert werden, die u.a. die Residualwerte für die jeweilige Ebene enthalten und für die Prüfung der Voraussetzungen mithilfe von SPSS genutzt werden können (Raudenbush et al., 2011).

a) Die Prüfung linearer Zusammenhänge zwischen Prädiktorvariablen und abhängiger Variable wird in der vorliegenden Untersuchung mittels Streudiagrammen analysiert. Dabei werden die Residualwerte der Prädiktoren auf Individual- und auf Klassenebene der jeweiligen abhängigen Variable gegenübergestellt und visuell geprüft, ob eine zufällige Streuung um den Wert null vorliegt (Raudenbush & Bryk, 2002).

b) Das Softwareprogramm HLM (Raudenbush et al., 2011) bietet für die Überprüfung homogener Fehlervarianzen einen integrierten Test an, der für die vorliegende Untersuchung verwendet wird. Eine Verletzung der Voraussetzungen ist häufig ein Hinweis dafür, dass das berechnete Modell noch nicht ausreichend gut spezifiziert ist (z.B. fehlende Berücksichtigung von relevanten Prädiktoren oder von Random Slopes; Raudenbush & Bryk, 2002). Zudem kann sich bei nicht normalverteilten Variablen mit stark positiver Kurtosis dieser Umstand als Fehlervarianzheterogenität zeigen, sodass Extremwerte ebenfalls dazu führen können, dass der entsprechende Test Varianzheterogenität anzeigt (Raudenbush & Bryk, 2002). Extremwerte bzw. Ausreisser sollten daher aus dem Datensatz entfernt werden (Raudenbush & Bryk, 2002).

c) Die Normalverteilung von Fehlertermen auf Individualebene kann visuell mit Q-Q-Diagrammen unter Einfügung einer normalverteilten Testverteilung sowie mit Histogrammen mit eingefügter Normalverteilungskurve in SPSS visuell untersucht werden und zusätzlich mittels der Faustregeln von Lienert und Raatz (1998) und Curran et al. (1996) beurteilt werden (vgl. Kap. 5.6.1).

d) Die Prüfung normalverteilter Fehlerterme auf Klassenebene wird in der vorliegenden Untersuchung mit der Mahalanobis-Distanz durchgeführt, wobei diese annähernd χ^2-verteilt sein sollte (Freiheitsgrade = Anzahl zufälliger Effekte + 1; Raudenbush & Bryk, 2002). Zusätzlich wird auch die empfohlene Überprüfung normalverteilter Fehlerterme auf Klassenebene mithilfe von Streudiagrammen verwendet, bei denen die Variablen CHIPCT[22] und die Variable MDIST (Mahalanobis-Distanz) einander gegenübergestellt werden, wobei die Werte auf einer 45-Grad-Linie liegen sollten (Raudenbush et al., 2011). Bei leichten Verletzungen dieser Voraussetzung kann auf die fixen Effekte mit robusten Standardfehlern zurückgegriffen werden, die in HLM ebenfalls standardmässig berechnet werden (Raudenbush & Bryk, 2002). Wenn sich zeigt, dass die robusten Standardfehler sich nicht wesentlich von den nicht robusten Standardfehlern unterscheiden, kann angenommen werden, dass die robusten Standardfehler trotz allenfalls gravierenderen Verletzungen der Normalverteilungsvoraussetzung valid geschätzt sind (Raudenbush & Bryk, 2002).

Tabelle 30: Prüfung der Voraussetzungen von Mehrebenenanalysen für die Mehrebenenmodelle der vorliegenden Arbeit

Voraussetzung	**Prüfung der Voraussetzung**
a) Lineare Zusammenhänge zwischen Prädiktorvariablen und abhängiger Variable	*Alle Modelle erfüllen die Anforderungen*
b) Homogene Varianzen der Fehlerterme (σ^2_e) auf Individualebene	*Alle Modelle erfüllen die Anforderungen unter folgender Bedingung:* - Ausschluss eines Kindes (Ausreisser) aus Stichprobe^{alle} in allen Modellen für den Follow-up 1 und Follow-up 2 (Modell 2, 3, 8a, 8b, 9a, 9b, 13 und 15) - Ausschluss eines Kindes (Ausreisser) aus der Stichprobe^{IQstreng} für alle Modelle beim Follow-up 1 und beim Follow-up 2 (Modell 5, 6, 11a, 11b, 12a. 12b, 14 und 16)
c) Normalverteilte Fehlerterme (e_{ij}) auf Individualebene	*Alle Modelle erfüllen die Anforderungen*
d) Normalverteilte Fehlerterme (u_{0j}) auf Klassenebene	χ^2-verteilte Mahalanobis-Distanz: *Keines der Modelle erfüllt die Anforderungen* Streudiagramm (CHIPCT vs. MDIST): *Alle Modelle erfüllen die Anforderungen mit Ausnahme der Modelle 12a und 12b (Stichprobe^{IQstreng}, Follow-up 2)*

[22] „Expected values of the order statistics for a sample of size *J* selected from a population that is distributed $\chi^2_{(v)}$“; Raudenbush, Bryk, Cheong, Congdon und Du Toit (2011, S. 40).

Die Prüfung der Voraussetzung für die Mehrebenenmodelle hat gezeigt, dass ein grosser Teil der Anforderungen erfüllt ist oder deren Verletzung durch zur Verfügung stehende Massnahmen entschärft werden kann (vgl. Tabelle 30). Bezüglich der Voraussetzung b) konnte die Verletzung durch den Ausschluss eines Kindes mit Extremwerten aus dem Datensatz behoben werden. Als Folge der Verletzung der Voraussetzung d) werden in der folgenden Arbeit zudem für alle Modelle robuste Standardfehler präsentiert. Da sich die robusten Standardfehler in allen Modellen dieser Arbeit nur unwesentlich von den nicht robusten Standardfehlern unterscheiden, kann angenommen werden, dass die robusten Standardfehler valid geschätzt sind (Raudenbush & Bryk, 2002).

Zusätzlich zu den genannten Voraussetzungen für Mehrebenenanalysen muss bei allen Regressionsanalysen auch geprüft werden, ob zwischen den Prädiktoren Multikollinearität vorliegt (H. Schneider, 2009). Da bei der Berechnung von Korrelationen mit SPSS die hierarchische Datenstruktur nicht berücksichtigt werden kann, werden in der vorliegenden Untersuchung die Korrelationen mithilfe von HLM (Raudenbush et al., 2013) berechnet. Dafür müssen die Variablen z-standardisiert, einer der Prädiktoren als abhängige Variable definiert und der andere Prädiktor als unabhängige Variable in das Mehrebenenmodell eingefügt werden. Der Regressionskoeffizient für die unabhängige Variable zeigt dann, wie hoch die beiden Prädiktoren unter Berücksichtigung der hierarchischen Datenstruktur miteinander korrelieren. Die Korrelationen zwischen den Prädiktoren finden sich im nächsten Kapitel (Kap. 6.1).

6. Ergebnisse

In diesem Kapitel werden die Ergebnisse der vorliegenden Untersuchung präsentiert. Zunächst werden die Korrelationen zwischen den Prädiktoren aufgezeigt sowie der ICC_1 für die Nullmodelle der einzelnen Messzeitpunkte genannt. Im Anschluss daran werden die Ergebnisse zur ersten Fragestellung dargestellt, gefolgt von den Ergebnissen für die zweite und die dritte Fragestellung. Wie sich noch zeigen wird, werfen die Ergebnisse jedoch Fragen auf. Daher werden auch weiterführende Mehrebenenmodelle berichtet und Analysen nur für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen durchgeführt, mit denen diesen offenen Fragen nachgegangen werden kann.

6.1 Korrelationen zwischen den Prädiktoren

Die folgenden Tabellen (Tabelle 31 bis Tabelle 36) informieren über die Korrelationen zwischen den verschiedenen Prädiktoren auf Individualebene für die Stichprobe$^{\text{alle}}$ und die Stichprobe$^{\text{IQstreng}}$ für den Post-Test, den Follow-up 1 und den Follow-up 2 unter Berücksichtigung der hierarchischen Datenstruktur. Es zeigt sich, dass bei allen Messzeitpunkten und somit in beiden Untersuchungsstichproben keine gravierenden Fälle von Multikollinearität anzunehmen sind (Backhaus, Erichson, Plinke & Weiber, 2016).

Stichprobe^{alle}

In der Stichprobe$^{\text{alle}}$ liegen beim Post-Test die höchsten Korrelationen zwischen mathematischem Vorwissen und IQ (r = .443, p < .001) und zwischen mathematischem Vorwissen und Förderbedarf in Mathematik (r = –.441, p < .001; vgl. Tabelle 31) vor. Der Zusammenhang zwischen IQ und Förderbedarf in Mathematik beträgt r = .324 (p < .001). Nicht signifikant sind die Korrelationen zwischen IQ und Geschlecht, SES und Geschlecht sowie Einschränkungen im Deutschverstehen und Geschlecht. Alle anderen Korrelationen sind ebenfalls signifikant, die Korrelationskoeffizienten sind jedoch kleiner r = .300 bzw. kleiner r = –.300 und können direkt der Tabelle 31 entnommen werden. Für die Stichprobe$^{\text{alle}}$ ergibt sich auf Basis des Nullmodells für den ICC_1 ein Wert von ρ = 0.149, sodass 14.9% der Varianz in der Mathematikleistung beim Post-Test durch Unterschiede zwischen den Klassen erklärt werden können und die Berechnung von Mehrebenenanalysen als sinnvoll erachtet werden kann (Muthén, 1997).

Tabelle 31: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe^{alle} beim Post-Test unter Berücksichtigung der hierarchischen Datenstruktur (N = 851)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.443***	–				
3. SES	.186***	.209***	–			
4. Alter	–.114**	–.248***	–.153***	–		
5. Geschlecht	.119***	–.011	–.034	.081*	–	
6. Förderbedarf Mathematik	–.441***	–.324***	–.126**	.098**	–.079**	–
7. Einschränkungen im Deutschverstehen	–.290***	–.168***	–.268***	.144**	.052	.164***

*** $p < .001$; ** $p < .010$; * $p < .050$

Beim Follow-up 1 ergibt sich für die Stichprobe^{alle} ein sehr ähnliches Bild wie beim Post-Test (vgl. Tabelle 32). Am stärksten korrelieren hier jedoch das mathematische Vorwissen und der Förderbedarf in Mathematik (r = –.540, p < .001). Zudem ist die Korrelation zwischen Alter und Geschlecht beim Follow-up 1 hochsignifikant (r = .081, p < .010). Bei allen anderen Korrelationen gibt es keine oder nur minimale Abweichungen (Differenzen < .01) im Vergleich zum Post-Test. Die Korrelationen für die restlichen Prädiktorvariablen können daher der Tabelle 32 entnommen werden. Der Wert für den ICC_1 beträgt für die Stichprobe^{alle} beim Follow-up 1 $\rho = 0.130$, sodass noch 13.0% der Varianz in der Mathematikleistung beim Follow-up 1 durch Unterschiede zwischen den Klassen erklärt werden können. Dieser Wert ist etwas niedriger als beim Post-Test, rechtfertigt aber nach wie vor die Anwendung von Mehrebenenanalysen (Muthén, 1997).

Tabelle 32: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe^{alle} beim Follow-up 1 unter Berücksichtigung der hierarchischen Datenstruktur (N = 856)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.442***	–				
3. SES	.184***	.203***	–			
4. Alter	–.112**	–.243***	–.158***	–		
5. Geschlecht	.110***	–.032	–.048	.081**	–	
6. Förderbedarf Mathematik	–.540***	–.321***	–.133**	.096**	–.073**	–
7. Einschränkungen im Deutschverstehen	–.290***	–.167***	–.275***	.146**	.057	.162***

*** $p < .001$; ** $p < .010$; * $p < .050$

Auch beim Follow-up 2 zeigt die Korrelationsmatrix in Tabelle 33, dass die Korrelationen zwischen den verschiedenen Individualprädiktoren in der Stichprobe^{alle} ähnlich ausfallen wie bisher beschrieben. Wie beim Follow-up 1 entfällt die höchste Korrelation auch beim Follow-up 2 auf den Zusammenhang zwischen mathematischem Vorwissen und Förderbedarf in Mathematik ($r = -.546$, $p < .001$). Zu nennen sind im Vergleich zum Post-Test zudem höhere Korrelationen für die Zusammenhänge zwischen mathematischem Vorwissen und Einschränkungen im Deutschverstehen ($r = -.302$, $p < .001$), zwischen IQ und Einschränkungen im Deutschverstehen ($r = -.180$, $p < .001$), zwischen SES und Förderbedarf in Mathematik ($r = -.153$, $p < .001$), zwischen SES und Einschränkungen im Deutschverstehen ($r = -.281$, $p < .001$), zwischen Alter und Einschränkungen im Deutschverstehen ($r = .160$, $p < .001$) und zwischen Förderbedarf in Mathematik und Einschränkungen im Deutschverstehen ($r = .182$, $p < .001$). Niedriger fällt hingegen die Korrelation zwischen Alter und Förderbedarf in Mathematik aus ($r = .084$, $p < .050$). Alle anderen Abweichungen sind sehr gering (Differenz < .01) und werden daher nicht weiter ausgeführt. Für die Stichprobe^{alle} ergibt sich für den ICC_1 ein Wert von $\rho = 0.119$, sodass 11.9% der Varianz in der Mathematikleistung beim Follow-up 2 durch Unterschiede zwischen den Klassen erklärt werden können, womit sich auch hier Mehrebenenanalysen empfehlen (Muthén, 1997).

Tabelle 33: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe^{alle} beim Follow-up 2 unter Berücksichtigung der hierarchischen Datenstruktur ($N = 818$)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.442***	–				
3. SES	.193***	.212***	–			
4. Alter	–.106**	–.248***	–.154***	–		
5. Geschlecht	.117***	–.031	–.047	.088**	–	
6. Förderbedarf Mathematik	–.546***	–.328**	–.153***	.084*	–.074**	–
7. Einschränkungen im Deutschverstehen	–.302***	–.180***	–.281***	.160***	.046	.182***

*** $p < .001$; ** $p < .010$; * $p < .050$

Stichprobe^{IQstreng}

Für die Stichprobe^{IQstreng} zeigt sich beim Post-Test, dass die höchsten Korrelationen auf die Zusammenhänge zwischen mathematischem Vorwissen und Förderbedarf in Mathematik ($r = -.528$, $p < .001$), zwischen mathematischem Vorwissen und IQ ($r = .445$, $p < .001$) und zwischen IQ und Förderbedarf in Mathematik entfallen ($r = -.314$, $p < .001$; vgl. Tabelle 34). Im Gegensatz zur Stichprobe^{alle} ergeben sich in der Stichprobe^{IQstreng} beim Post-Test keine signifikanten Korrelationen zwischen Alter und mathematischem Vorwissen ($r = -.054$, $p = .173$), Alter und Geschlecht ($r = .038$, $p = .193$), Alter und Förderbedarf in Mathematik ($r = .033$, $p = .397$) sowie Alter und Einschränkungen im Deutschverstehen ($r = .076$, $p = .076$). Dies kann dadurch erklärt werden, dass in der Stichprobe^{IQstreng} Kinder, älter als 9 Jahre und 5 Monate, aufgrund der verwendeten IQ-Altersnorm von den Analysen ausgeschlossen worden sind. Diese Kinder weisen im Vergleich zu den Kindern, die die Altersnorm erfüllen, im Durchschnitt weniger mathematisches Vorwissen auf, sind häufiger Knaben und haben eher einen Förderbedarf in Mathematik oder Einschränkungen im Deutschverstehen (vgl. Tabelle 23). Signifikant ist in Stichprobe^{IQstreng} beim Post-Test hingegen der Zusammenhang zwischen Geschlecht und Förderbedarf in Mathematik ($r = -.081$, $p < .010$). Weiter ist zu beachten, dass in der Stichprobe^{IQstreng} die Korrelation zwischen SES und Alter ($r = -.077$, $p < .050$) deutlich geringer ist als in der Stichprobe^{alle} und nur noch auf dem 5%-Niveau signifikant ist. Auch dies lässt sich durch den Ausschluss der oben genannten Gruppe von Kindern erklären. Alle weiteren Korrelationen können der Tabelle 34 entnommen werden. Der ICC_1 liegt beim Post-Test für die Stichprobe^{IQstreng} bei ρ = 0.165, was bedeutet, dass 16.5% der Varianz der Mathematikleistung durch Unterschiede zwischen den Klassen erklärt werden können. Auch für diese Untersuchungsstichprobe sind daher Mehrebenenanalysen angezeigt (Muthén, 1997).

Tabelle 34: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe^{IQstreng} beim Post-Test unter Berücksichtigung der hierarchischen Datenstruktur ($N = 806$)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.445***	–				
3. SES	.175***	.194***	–			
4. Alter	–.054	–.211***	–.077*	–		
5. Geschlecht	.110***	.003	–.025	.038	–	
6. Förderbedarf Mathematik	–.528***	–.314***	–.114**	.033	–.081**	–
7. Einschränkungen im Deutschverstehen	–.263***	–.161***	–.250***	.076	.070*	.150***

*** $p < .001$; ** $p < .010$; * $p < .050$

Beim Follow-up 1 sind die Korrelationen zwischen den Individualprädiktoren für die Stichprobe^{IQstreng} (fast) identisch mit den Werten beim Post-Test für diese Untersuchungsstichprobe (Differenzen < .01). Die einzelnen Werte werden daher an dieser Stelle nicht weiter ausgeführt (vgl. Tabelle 35). Der ICC_1 liegt für die Stichprobe^{IQstreng} beim Follow-up 1 bei ρ = 0.149, sodass 14.9% der Varianz in der Mathematikleistung durch Unterschiede zwischen den Klassen erklärt werden können.

Tabelle 35: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe$^{\text{IQstreng}}$ beim Follow-up 1 unter Berücksichtigung der hierarchischen Datenstruktur (N = 811)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.443***	–				
3. SES	.173***	.187***	–			
4. Alter	–.049	–.204***	–.083*	–		
5. Geschlecht	.101***	–.019	–.040	.038	–	
6. Förderbedarf Mathematik	–.527***	–.310***	–.121**	.029	–.075**	–
7. Einschränkungen im Deutschverstehen	–.263***	–.159***	–.257***	.077	.076*	.148***

*** $p < .001$; ** $p < .010$; * $p < .050$

Auch beim Follow-up 2 zeigt sich ein ähnliches Bild für die Korrelationen der Individualprädiktoren in der Stichprobe$^{\text{IQstreng}}$ im Vergleich zum Post-Test (vgl. Tabelle 36). Allerdings sind die Korrelationen zwischen SES und Geschlecht (r = –.038, p = .208) und zwischen Geschlecht und Einschränkungen im Deutschverstehen (r = .059, p = .084) beim Follow-up 2 nicht signifikant. Auch fallen die Korrelationen zwischen SES und Förderbedarf in Mathematik (r = –.138, p < .010), zwischen SES und Einschränkungen im Deutschverstehen (r = –.260, p < .001) sowie zwischen Förderbedarf in Mathematik und Einschränkungen im Deutschverstehen (r = .163, p < .001) etwas höher aus als beim Post-Test. Die Korrelation zwischen Geschlecht und Förderbedarf in Mathematik ist etwas niedriger und nur auf dem 5%-Niveau signifikant (r = –.067, p < .050). Beim Follow-up 2 liegt der ICC_1 in der Stichprobe$^{\text{IQstreng}}$ bei ρ = 0.124, sodass 12.4% der Varianz in der Mathematikleistung durch Unterschiede zwischen den Klassen erklärt werden können und Mehrebenenanalysen angezeigt sind (Muthén, 1997; Thompson, 2002).

Tabelle 36: Korrelationen zwischen den Prädiktoren auf Individualebene für die Stichprobe$^{\text{IQstreng}}$ beim Follow-up 2 unter Berücksichtigung der hierarchischen Datenstruktur (N = 778)

Variable	1	2	3	4	5	6
1. Mathematisches Vorwissen	–					
2. IQ „alle“	.445***	–				
3. SES	.180***	.198***	–			
4. Alter	–.006	–.210***	–.064	–		
5. Geschlecht	.104***	–.022	–.038	.056	–	
6. Förderbedarf Mathematik	–.536***	–.322***	–.138**	.034	–.067*	–
7. Einschränkungen im Deutschverstehen	–.271***	–.166***	–.260***	.085	.059	.163***

*** $p < .001$; ** $p < .010$; * $p < .050$

6.2 Prädiktoren für die Mathematikleistung auf Individual- und Klassenebene

In diesem Kapitel wird die Frage beantwortet, welche Prädiktoren auf Individual- und auf Klassenebene einen signifikanten Einfluss auf die abhängige Variable „Mathematikleistung“ zum jeweiligen Messzeitpunkt haben. Zuerst werden die Modelle für die Stichprobe^{alle} betrachtet. In einem zweiten Schritt wird untersucht, ob sich für die Stichprobe^{IQstreng} abweichende Modelle ergeben. Diese Stichprobe weist verlässlichere IQ-Schätzungen auf, ist jedoch weniger repräsentativ mit Bezug auf ältere Kinder (vgl. Kapitel 5.5). Auf eine Darstellung des schrittweisen Modellaufbaus wird in dieser Arbeit verzichtet. Bei den präsentierten Modellen handelt es sich um die finalen Modelle inklusive allfälliger Cross-Level-Interaktionen, die sich basierend auf den Devianzwerten am besten zu den Daten passend erwiesen haben (vgl. Kap. 5.6.2).

6.2.1 Stichprobe^{alle}

Stichprobe^{alle}: Post-Test

Für die Stichprobe^{alle} passt beim Post-Test im dritten Quartal der dritten Klasse das Modell 1 (vgl. Tabelle 37) am besten zu den Daten. Gemäss diesem Modell liegt die durchschnittliche Mathematikleistung aller Kinder aus der Stichprobe^{alle} bei 25.68 Punkten (SE = 0.40, p = .000). Auf *Individualebene* hat das mathematische Vorwissen beim Post-Test einen signifikanten Einfluss auf die Mathematikleistung bei einem Regressionskoeffizienten[23] von 0.81 (SE = 0.05, p = .000). Ein um einen Punkt höheres mathematisches Vorwissen im Vergleich zum durchschnittlichen Vorwissen der Stichprobe führt somit zu einer Mathematikleistung, die um 0.81 Punkte höher liegt (25.86 + 0.81 = 26.49). Der standardisierte Regressionskoeffizient β für das mathematische Vorwissen beträgt 0.57. Auch der IQ hat einen signifikanten Einfluss auf die Mathematikleistung beim Post-Test bei einem Regressionskoeffizienten von 0.08 (SE = 0.01, p = .000). Bei einem um einen Punkt höheren IQ-Wert verglichen mit dem durchschnittlichen IQ der Stichprobe, liegt somit eine um 0.08 Punkte höhere Mathematikleistung vor als der Wert für die durchschnittliche Mathematikleistung (25.68 + 0.08 = 26.76). Der standardisierte Regressionskoeffizient β für den IQ beträgt 0.12 und ist damit niedriger als der Wert für das mathematische Vorwissen. Auch das Geschlecht ist ein signifikanter Prädiktor für die Mathematikleistung beim Post-Test. Knaben weisen im Durchschnitt eine um 1.31 Punkte höhere Mathematikleistung auf als Mädchen (SE = 0.39, p = .001). Ein Förderbedarf in Mathematik wirkt sich beim Post-Test bei einem Regressionskoeffizienten von –4.03 (SE = 0.62, p = .000) negativ auf die Mathematikleistung aus und geht mit einer Mathematikleistung einher, die um –4.03 Punkte niedriger ist als der durchschnittliche Wert für die Stichprobe^{alle} (25.68 – 4.03 = 21.65).

[23] Es handelt sich hierbei um unstandardisierte Regressionskoeffizienten (*B*). Für gewisse Variablen wurden auch standardisierte Regressionskoeffizienten (β) berechnet (vgl. Kap. 5.6.2), und explizit als solche bezeichnet.

Tabelle 37: Mehrebenenmodelle für die Mathematikleistung beim Post-Test, Follow-up 1 und Follow-up 2 mit Individual- und Klassenprädiktoren für die Stichprobe[alle]

	Modell 1: Post-Test			**Modell 2: Follow-up 1**			**Modell 3: Follow-up 2**		
	B	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***
Intercept	**25.68**	0.40	.000	**27.82**	0.34	.000	**44.50**	0.49	.000
Individualebene									
Vorwissen Mathematik	**0.81**	0.05	.000	**0.82**	0.04	.000	**0.94**	0.07	.000
IQ	**0.08**	0.01	.000	**0.05**	0.01	.001	**0.10**	0.02	.000
Alter (in Monaten)				**–0.09**	0.03	.007			
Geschlecht[1)]	**1.31**	0.39	.001	**1.50**	0.30	.000	**0.92**	0.43	.032
Förderbedarf Mathematik[2)]	**–4.03**	0.62	.000	**–4.20**	0.60	.000	**–3.57**	0.88	.000
Einschränkungen im Deutschverstehen[3)]							–0.79	0.54	.147
Klassenebene									
IQ-Mittelwert der Klassen	**0.15**	0.06	.011	**0.12**	0.05	.021			
Anteil Knaben in den Klassen (%)				**0.09**	0.03	.001			
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (%)				**–0.04**	0.02	.039			
Cross-Level									
Förderbedarf Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen							**–0.10**	0.04	.005
	Var	***SD***	***p***	***Var***	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte									
σ^2_e (Fehlervarianz L1)	26.64	5.16		22.64	4.76		36.53	6.04	
σ^2_{u0} (Fehlervarianz L2)	4.00	2.00	.000	2.93	1.71	.000	6.72	2.59	.000
σ^2_{u1} (Vorwissen Mathematik)							0.08	**0.28**	.003
σ^2_{u4} (Förderbedarf Mathematik)							17.17	**4.14**	.001
Erklärte Varianz L1 (%)[4)]		62.22			65.91			57.89[5)]	
Erklärte Varianz L2 (%)[4)]		15.25			14.36			10.69[5)]	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; [5)] Varianzaufklärung basiert auf Modell ohne Random Slopes: Varianzkomponente L1 = 39.18, Varianzkomponente L2 = 6.20; *Var* = Varianzkomponente

Auf *Klassenebene* hat in der Stichprobe^{alle} nur der IQ-Mittelwert der Klasse bei einem Regressionskoeffizienten von 0.15 (*SE* = 0.06, *p* = .011) einen signifikanten Einfluss auf die Mathematikleistung beim Post-Test. Demnach erhöht sich bei einer Klasse, deren IQ-Mittelwert um einen Punkt höher liegt als der durchschnittliche IQ-Wert aller untersuchten Klassen, die Mathematikleistung um 0.15 Punkte (25.68 + 0.15 = 25.83).[24] Der standardisierte Regressionskoeffizient β liegt für den IQ-Mittelwert der Klasse bei 0.10 und ist somit vergleichbar hoch wie der Wert für den individuellen IQ. Das Modell 1 für die Stichprobe^{alle} erklärt beim Post-Test auf Individualebene 62.22% der Varianz und auf Klassenebene 15.25% der Varianz (vgl. Tabelle 37).

Stichprobe^{alle}: Follow-up 1

Beim Follow-up 1 hat sich für die Stichprobe^{alle} das Modell 2 als das am besten zu den Daten passende Modell erwiesen (vgl. Tabelle 37). Die durchschnittliche Mathematikleistung beträgt beim Follow up 1 für diese Untersuchungsstichprobe unter Berücksichtigung der im Modell 2 genannten Prädiktoren demnach 27.82 Punkte (*SE* = 0.34, *p* = .000), sodass insgesamt ein Lernzuwachs im Vergleich zum Post-Test festgestellt werden kann.[25]

Auf *Individualebene* hat das mathematische Vorwissen mit einem Regressionskoeffizienten von 0.82 (*SE* = 0.04, *p* = .000) ebenso wie der IQ mit einem Regressionskoeffizienten von 0.05 (*SE* = 0.01, *p* = .001) einen signifikanten Einfluss auf die Mathematikleistung beim Follow-up 1. Der standardisierte Regressionskoeffizient liegt für das mathematische Vorwissen bei 0.60 und für den IQ bei 0.08. Der Wert für das Vorwissen bewegt sich in einer ähnlichen Grössenordnung wie der Wert beim Post-Test, der standardisierte Regressionskoeffizient für den IQ ist jedoch einiges niedriger. Zudem zeigt sich beim Follow-up 1 im Gegensatz zum Post-Test, dass das Alter ebenfalls in einem signifikanten Zusammenhang mit der Mathematikleistung steht. Bei einem Kind, dessen Alter um einen Monat höher ist als der Durchschnittswert für die Stichprobe, liegt somit eine um 0.09 Punkte niedrigere Mathematikleistung vor (*B* = – 0.09, *SE* = 0.03, *p* = .007). Weiter bestätigt sich erneut, dass Knaben eine höhere Mathematikleistung (*B* = 1.50, *SE* = 0.30, *p* = .000) und Kinder mit einem Förderbedarf in Mathematik eine niedrigere Mathematikleistung erreichen (*B* = –4.20, *SE* = 0.60, *p* = .000).

Auf *Klassenebene* wiederum zeigt sich, dass in Klassen, die im Vergleich zum Durchschnittswert aller Klassen einen um einen Punkt höheren IQ-Mittelwert haben, die Mathematikleistung

[24] Mithilfe der jeweiligen Regressionskoeffizienten kann bei den anderen Modellen in gleicher Weise die effektive Mathematikleistung unter Berücksichtigung der interessierenden Prädiktorvariablen berechnet werden.
[25] Da beim Post-Test und beim Follow-up 1 der gleiche Mathematiktest eingesetzt wurde, können die Intercepts miteinander verglichen werden (vgl. Kapitel 5.3.2).

beim Follow-up 1 um 0.12 Punkte ansteigt (*SE* = 0.05, *p* = .021), wobei der standardisierte Regressionskoeffizient fast identisch ist mit dem Wert beim Post-Test (β = 0.09). Zusätzlich hat beim Follow-up 1 im Gegensatz zum Post-Test auch der Anteil Knaben und Mädchen in der Klasse einen signifikanten Einfluss auf die abhängige Variable. Demnach resultiert in Klassen, in denen im Vergleich zum Durchschnittswert für alle Klassen ein Prozent mehr Knaben die Klasse besuchen, eine um 0.09 Punkte höhere Mathematikleistung (*SE* = 0.03, *p* = .001). Zudem wirkt sich neu beim Follow-up 1 auch ein im Vergleich zur gesamten Stichprobe erhöhter Anteil an Kindern mit Einschränkungen im Deutschverstehen bei einem Regressionskoeffizienten von –0.04 signifikant auf die Mathematikleistung aus (*SE* = 0.02, *p* = .039). Beim Follow-up 1 kann das erstellte Modell auf Individualebene 65.91% der Varianz und auf Klassenebene 14.36% der Varianz erklären. Demnach wird in der Stichprobe^{alle} im Vergleich zum Modell 1 beim Post-Test im Modell 2 für den Follow-up 1 auf Individualebene mehr, auf Klassenebene hingegen weniger Varianz erklärt.

Stichprobe^{alle}: Follow-up 2

Für die Stichprobe^{alle} liegt der Intercept für den Follow-up 2 am Ende der vierten Klasse bei 44.50 Punkten[26] (*SE* = 0.49, *p* = .000; vgl. Modell 3 in Tabelle 37). Auf *Individualebene* hat das mathematische Vorwissen (*B* = 0.94, *SE* = 0.07, *p* = .000) in der Stichprobe^{alle} auch beim Follow-up 2 einen signifikanten Einfluss auf die Mathematikleistung, wobei der standardisierte Regressionskoeffizient identisch ist mit dem Wert beim Post-Test (β = 0.57). Ebenfalls signifikant ist der Einfluss des IQs (*B* = 0.10, *SE* = 0.02, *p* = .000), der standardisierte Regressionskoeffizient ist hier leicht höher als bei den vorangegangenen Messzeitpunkten (β = 0.14). Für das mathematische Vorwissen zeigt sich jedoch beim Follow-up 2, dass der Zusammenhang zwischen dem mathematischen Vorwissen und der Mathematikleistung nicht in allen Klassen gleich ausfällt (*SD* = 0.28, *p* = .003; vgl. zufällige Effekte). Mithilfe der Standardabweichung der Random Slope für das mathematische Vorwissen kann ein 95%-Konfidenzintervall berechnet werden ($0.94 \pm 0.28 \times 1.96$; vgl. Hox, 2010, S. 19). Demnach liegen in 95% aller Klassen die Regressionskoeffizienten für das mathematische Vorwissen zwischen 0.39 und 1.49. In einzelnen Klassen ist der Einfluss des Vorwissens somit knapp viermal so gross wie in Klassen mit dem geringsten Einfluss des Vorwissens. Es konnte keine Klassenvariable identifiziert werden, die diesen Unterschied erklären kann. Beim Geschlecht zeigt sich auch beim Follow-up 2, dass Knaben im Durchschnitt eine höhere Mathematikleistung aufweisen (*B* = 0.92, *SE* =

[26] Die Mathematikleistungen beim Follow-up 2 lassen sich nicht direkt mit den Mathematikleistungen beim Post-Test und beim Follow-up 1 vergleichen, da unterschiedliche Tests mit verschiedener Anzahl Aufgaben verwendet wurden (vgl. Kapitel 5.3.2).

0.43, p = .032). Für den Förderbedarf in Mathematik beträgt der Regressionskoeffizient –3.57 Punkte *(SE* = 0.88, p = .000), allerdings müssen auch hier Random Slopes berücksichtigt werden (*SD* = 4.14, p = .001). Demnach liegt der Regressionskoeffizient für 95% aller Klassen zwischen –11.68 und 4.54 (–3.57 ± 4.14 × 1.96). Das bedeutet, dass sich in gewissen Klassen ein Förderbedarf in Mathematik stark negativ auf die Mathematikleistung, in anderen Klassen jedoch sogar positiv auf die Mathematikleistung beim Follow-up 2 ausgewirkt hat.

Die *Cross-Level-Interaktion* „Förderbedarf in Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen" besagt zusätzlich, dass der Zusammenhang zwischen Förderbedarf in Mathematik und Mathematikleistung beim Follow-up 2 teilweise durch den Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen erklärt werden kann (*B* = –0.10, *SE* = 0.04, p = .005). Demnach führt ein um ein Prozent höherer Anteil an Kindern mit Einschränkungen im Deutschverstehen im Vergleich zum Durchschnittswert der Untersuchungsstichprobe dazu, dass sich der negative Zusammenhang zwischen Förderbedarf in Mathematik und Mathematikleistung beim Follow-up 2 um –0.10 Punkte verstärkt. D.h., der Regressionskoeffizient für den Förderbedarf in Mathematik beträgt für eine solche Klasse ohne Berücksichtigung der Random Slope: –3.83 – 0.10 = –3.93. Die Abbildung 10 zeigt grafisch den unterschiedlichen Einfluss des Anteils Kinder mit Einschränkungen im Deutschverstehen in den Klassen auf den Zusammenhang zwischen dem Förderbedarf in Mathematik und der Mathematikleistung beim Follow-up 2.

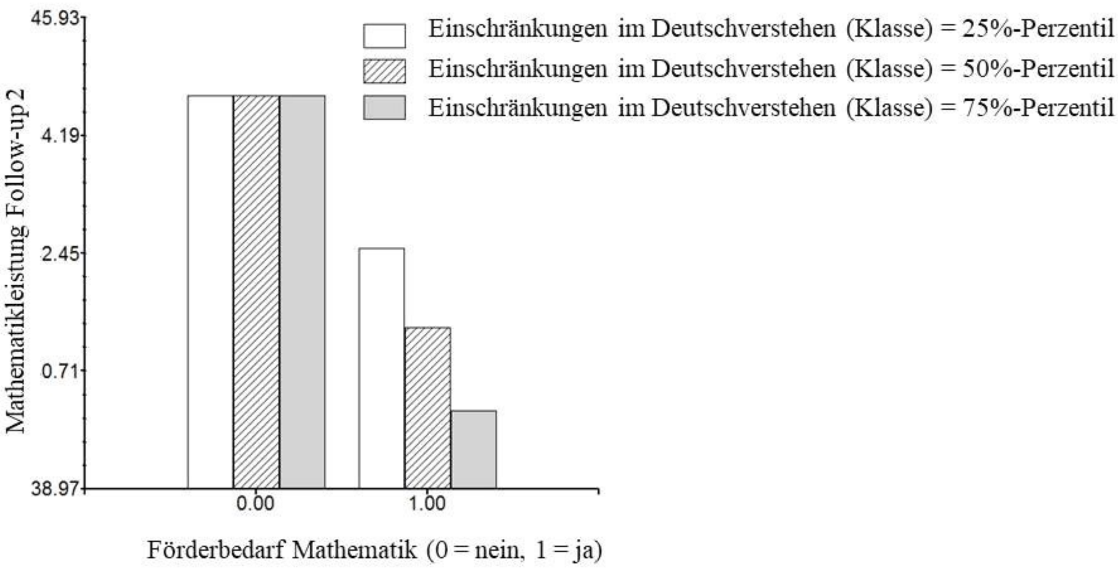


Abbildung 10: Zusammenhang zwischen Mathematikleistung beim Follow-up 2 und Förderbedarf in Mathematik in Abhängigkeit des Anteils Kinder mit Einschränkungen im Deutschverstehen in den Klassen für die Stichprobe^{alle}

Einschränkungen im Deutschverstehen haben als Individualprädiktor unter Berücksichtigung der bereits genannten Prädiktoren keinen signifikanten Einfluss auf die Mathematikleistung beim Follow-up 2 (B = –0.79, SE = 0.54, p = .147). Der Prädiktor wurde dennoch im Modell belassen, da sich bei der schrittweisen Modellentwicklung gezeigt hat, dass es sich hierbei um einen signifikanten Prädiktor handelt und ein Hinzufügen zu einer signifikanten Modellverbesserung führt (Hox, 2010). Erst durch das Hinzufügen der Random Slopes hat sich das Signifikanzniveau dieses Prädiktors verändert. Die Varianzaufklärung für das Modell 3 beim Follow-up 2 beträgt für die Stichprobe^{alle} auf Individualebene 57.89% und auf Klassenebene 10.69%.[27] Diese Werte sind somit auf beiden Ebenen kleiner als für die bisher berichteten Modelle.

Zusammenfassend zeigen die präsentierten Modelle für die Stichprobe^{alle}, dass das mathematische Vorwissen, der IQ, das Geschlecht und ein Förderbedarf in Mathematik bei allen drei Messzeitpunkten auf *Individualebene* signifikante Prädiktoren für die Mathematikleistung sind. Demnach weisen Kinder, die über viel Vorwissen bzw. einen hohen IQ beim Pre-Test verfügen, beim Post-Test und bei den beiden Follow-up-Messungen höhere Mathematikleistungen auf. Ein Vorhandensein eines Förderbedarfs in Mathematik hingegen geht mit einer deutlich niedrigeren Mathematikleistung einher. Beim Geschlecht wiederum zeigen die Modelle, dass Knaben im Durchschnitt höhere Mathematikleistungen erreichen als Mädchen. Zu den Prädiktoren auf Individualebene muss jedoch angemerkt werden, dass der Zusammenhang zwischen mathematischem Vorwissen und Mathematikleistung sowie zwischen einem Förderbedarf in Mathematik und der Mathematikleistung beim Follow-up 2 nicht für alle Kinder gleich stark ausfällt (Random Slopes). Auf *Klassenebene* hingegen gibt es in der Stichprobe^{alle} keinen Prädiktor, der sich bei allen Messzeitpunkten als signifikant erwiesen hat. Beim Post-Test und beim Follow-up 1 ergeben die Modelle, dass sich ein hohes Intelligenzniveau der Klasse positiv auf die individuellen Mathematikleistungen auswirkt, nicht jedoch beim Follow-up 2. Zusätzlich hat beim Follow-up 1 auch ein über dem Durchschnitt liegender Anteil an Knaben in den Klassen einen positiven Einfluss auf die Mathematikleistung. Liegt der Anteil an Kindern mit Einschränkungen im Deutschverstehen in einer Klasse zudem über dem Durchschnittswert für die Stichprobe, hat dies einen negativen Einfluss auf die individuelle Mathematikleistung beim Follow-up 1. Beim Follow-up 2 zeigt sich auf Klassenebene einzig, dass sich der Anteil an Kindern mit Einschränkungen im Deutschverstehen in den Klassen signifikant auf den Zusammenhang zwischen dem Förderbedarf in Mathematik und der Mathematikleistung auswirkt

[27] Da die Berechnung der Varianzaufklärung gemäss Korrekturvorschlag von Snijders und Bosker (1994) bei Random Slopes sehr komplex ist, basieren die genannten Werte auf einem Modell, in dem die gleichen Prädiktoren wie im Modell 3 berücksichtigt sind, jedoch die Random Slopes weggelassen wurden (vgl. auch Tabelle 37).

(*Cross-Level-Interaktion*), sodass in Klassen mit einem hohen Anteil an Kindern mit Einschränkungen im Deutsch Verstehen sich ein Förderbedarf in Mathematik stärker negativ auf die Mathematikleistung beim Follow-up 2 auswirkt.

6.2.2 Stichprobe$^{\text{IQstreng}}$

Stichprobe^{IQstreng}: Post-Test

In der Stichprobe$^{\text{IQstreng}}$ ergibt sich für den Post-Test auf Klassenebene kein eindeutiges Bild. Im Modell 4a zeigt sich, dass das mathematische Vorwissen der Klasse (B = 0.31, SE = 0.16, p = .048) knapp ein signifikanter Prädiktor auf *Klassenebene* ist (vgl. Tabelle 38).[28] Im Modell 4b hingegen erweist sich der IQ-Mittelwert der Klasse als signifikanter Prädiktor (B = 0.15, SE = 0.05, p = .011). Beide Modelle haben sich als signifikant besser als das Modell ohne jegliche Prädiktoren auf Klassenebene erwiesen. Aufgrund der Devianzstatistik kann jedoch nicht entschieden werden, welches Modell besser zu den Daten passt, da hierfür eine unterschiedliche Anzahl Prädiktoren in den zu vergleichenden Modellen notwendig ist (Raudenbush & Bryk, 2002). Werden hingegen beide Prädiktoren gemeinsam berücksichtigt (vgl. Modell 4c in Tabelle 38), ist dieses Modell weder besser als das Modell 4a ($\chi^2(1, N_{L1} = 806, N_{L2} = 58) = 2.76, p$ = .093) noch besser als das Modell 4b ($\chi^2(1, N_{L1} = 806, N_{L2} = 58) = 1.58, p = .208$). Zudem ist unter gemeinsamer Berücksichtigung der beiden Prädiktoren keiner der beiden Prädiktoren signifikant (Vorwissen Mathematik der Klasse: B = 0.19, SE = 0.17, p = .265; IQ-Mittelwert der Klasse: B = 0.11, SE = 0.06, p = .087). Daraus lässt sich schliessen, dass das Vorwissen und der IQ-Mittelwert der Klassen einen gemeinsamen Anteil der Varianz in den Mathematikleistungen in der Stichprobe$^{\text{IQstreng}}$ erklären können und sich diese dabei so konkurrenzieren, dass bei gemeinsamer Berücksichtigung für beide Variablen kein signifikanter Effekt mehr vorliegt. Auch wenn nicht abschliessend geklärt werden kann, ob das Modell 4a oder das Modell 4b besser zu den Daten passt, zeigt sich auf Individualebene in beiden Modellen ein fast identisches Bild. Daher wird im Folgenden nur das Modell 4b auf Individualebene näher erläutert, da dieses bezüglich der berücksichtigten Prädiktoren stärker mit dem Modell 1 der Stichprobe$^{\text{alle}}$ für den Post-Test übereinstimmt und eine höhere Varianzaufklärung auf beiden Ebenen aufweist.

[28] Der standardisierte Regressionskoeffizient für das mathematische Vorwissen der Klasse beträgt 0.09.

Tabelle 38: Mehrebenenmodelle für die Mathematikleistung beim Post-Test mit Individual- und Klassenprädiktoren für die Stichprobe$^{\text{IQstreng}}$

	Modell 4a			**Modell 4b**			**Modell 4c**		
	B	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***
Intercept	**26.01**	0.41	.000	**26.05**	0.41	.000	**26.05**	0.40	.000
Individualebene									
Vorwissen Mathematik	**0.79**	0.05	.000	**0.80**	0.05	.000	**0.79**	0.05	.000
IQ	**0.08**	0.01	.000	**0.08**	0.01	.000	**0.08**	0.01	.000
Geschlecht[1)]	**1.36**	0.39	.001	**1.34**	0.39	.001	**1.35**	0.39	.001
Förderbedarf Mathematik[2)]	**–3.85**	0.60	.000	**-3.87**	0.60	.000	**–3.90**	0.59	.000
Einschränkungen im Deutschverstehen[3)]	**–0.87**	0.41	.036	**-0.88**	0.42	0.35	**–0.86**	0.42	.039
Klassenebene									
Vorwissen Mathematik der Klassen	**0.31**	0.16	.048				0.19	0.17	.265
IQ-Mittelwert der Klassen				**0.15**	0.05	.011	0.11	0.06	.087
	Var	***SD***	***p***	***Var***	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte									
σ^2_e (Fehlervarianz L1)	25.83	5.08		25.84	5.08		25.83	5.08	
σ^2_{u0} (Fehlervarianz L2)	4.17	2.04	.000	4.03	2.01	.000	4.00	2.00	.000
Erklärte Varianz L1 (%)[4)]		62.18			62.35			62.40	
Erklärte Varianz L2 (%)[4)]		16.85			17.05			17.10	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Im Modell 4b mit dem IQ-Mittelwert der Klasse als Prädiktor auf Klassenebene liegt die durchschnittliche Mathematikleistung in der Stichprobe$^{\text{IQstreng}}$ unter Berücksichtigung der im Modell genannten Prädiktoren bei 26.05 Punkten ($SE = 0.41$, $p = .000$) und ist somit um 0.37 Punkte höher als in der Stichprobe$^{\text{alle}}$ beim Post-Test (vgl. Modell 1 in Tabelle 37). Dies kann durch den Ausschluss von Kindern in der Stichprobe$^{\text{IQstreng}}$ erklärt werden, deren Alter höher ist als in der IQ-Altersnorm vorgesehen. Da diese Kinder über weniger mathematisches Vorwissen im Vergleich zur Gesamtstichprobe verfügen, wirkt sich dies erwartungskonform negativ auf die durchschnittliche Post-Test-Leistung im Mathematiktest in der Stichprobe$^{\text{alle}}$ aus (vgl. Kapitel 5.5).

Nur minimale Unterschiede bezüglich der Höhe der Regressionskoeffizienten im Vergleich zur Stichprobe$^{\text{alle}}$ ergeben sich auf *Individualebene* beim Post-Test für die Stichprobe$^{\text{IQstreng}}$ beim mathematischen Vorwissen (B = 0.80, SE = 0.05, p = .000), beim IQ (B = 0.08, SE = 0.01, p = .000) und beim Geschlecht (B = 1.34, SE = 0.39, p = .001). Der standardisierte Regressionskoeffizient für das mathematische Vorwissen beträgt 0.55 und für den IQ 0.13. Im Gegensatz zur Stichprobe^{alle} sind jedoch in der Stichprobe^{IQstreng} auch Einschränkungen im Deutschverstehen bei einem Regressionskoeffizienten von –0.87 (SE = 0.60, p = .000) auf Individualebene ein signifikanter Prädiktor für die Mathematikleistung beim Post-Test. Zudem fällt in der Stichprobe^{IQstreng} unter zusätzlicher Berücksichtigung von Einschränkungen im Deutschverstehen der Regressionskoeffizient für den Förderbedarf in Mathematik niedriger aus (B = –3.85, SE = 0.60, p = .000). Auf *Klassenebene* ist der Regressionskoeffizient für den IQ-Mittelwert der Klasse (B = 0.15, SE = 0.05, p = .011) in der Stichprobe^{IQstreng} hingegen gleich hoch wie in der Stichprobe^{alle}. Der standardisierte Regressionskoeffizient für den IQ-Mittelwert liegt bei 0.11. Das Modell 5 erklärt auf Individualebene 62.35% der Varianz und auf Klassenebene 17.05% der Varianz. Damit liegt der Wert auf Individualebene sehr nahe beim Wert für die Stichprobe^{alle}, auf Klassenebene ist dieser Wert hingegen etwas höher.

Stichprobe^{IQstreng}: Follow-up 1

Beim Follow-up 1 zeigt sich, dass für die Stichprobe^{IQstreng} ein Modell mit den gleichen Prädiktoren wie in der Stichprobe^{alle} am besten zu den Daten passt (vgl. Modell 5 in Tabelle 39). Der Intercept liegt beim Follow-up 1 für die Stichprobe^{IQstreng} bei 28.11 Punkten (SE = 0.34, p = .000), sodass der Unterschied zur Stichprobe^{alle} bei diesem Messzeitpunkt noch 0.29 Punkte beträgt (vgl. Modell 2 in Tabelle 37). Bezüglich der Regressionskoeffizienten und der Signifikanzniveaus ergeben sich nur minimale Unterschiede im Vergleich zum Modell 2 für die Stichprobe^{alle}.

Leichte Abweichungen auf *Individualebene* gibt es mit Bezug auf die Regressionskoeffizienten des IQs (B = 0.04, SE = 0.01, p = .002; β = 0.07) und des Geschlechts (B = 1.40, SE = 0.31, p = .000). Der standardisierte Regressionskoeffizient für den IQ ist mit 0.07 nur noch fast halb so hoch wie beim Post-Test, für das mathematische Vorwissen sind die Abweichungen jedoch gering (β = 0.60). In der Stichprobe^{IQstreng} zeigt sich beim Follow-up 1 ebenfalls ein negativer Zusammenhang zwischen Alter und Mathematikleistung, allerdings ist die Fehlerwahrscheinlichkeit, dass es sich hier um einen zufälligen Effekt handelt, grösser (B = –0.09, SE = 0.04, p = .023).

Auf *Klassenebene* ist zudem der Regressionskoeffizient für den IQ-Mittelwert der Klasse minimal höher bei einer geringeren Irrtumswahrscheinlichkeit (B = 0.13, SE = 0.05, p = .014) als in der Stichprobe^{alle}. Der standardisierte Regressionskoeffizient ist mit 0.10 fast identisch mit dem Wert für den Post-Test der Stichprobe^{IQstreng}. Der Regressionskoeffizient für den Anteil Kinder mit Einschränkungen im Deutschverstehen wiederum fällt gleich hoch aus, die Irrtumswahrscheinlichkeit ist jedoch niedriger (B = –0.04, SE = 0.02, p = .021). Das Modell 5 erklärt auf Individualebene 66.05% der Varianz und auf Klassenebene 16.71% der Varianz. Somit ist auch beim Follow-up 1 in der Stichprobe^{IQstreng} die Varianzaufklärung auf Individualebene sehr nahe beim Wert für die Stichprobe^{alle}, auf Klassenebene liegt dieser Wert hingegen etwas mehr als 2% höher.

Tabelle 39: Mehrebenenmodell für die Mathematikleistung beim Follow-up 1 mit Individual- und Klassenprädiktoren für die Stichprobe^{IQstreng}

	Modell 5		
	B	***SE***	***p***
Intercept	**28.11**	0.34	.000
Individualebene			
Vorwissen Mathematik	**0.82**	0.04	.000
IQ	**0.04**	0.01	.002
Alter (in Monaten)	**–0.09**	0.04	.023
Geschlecht[1)]	**1.40**	0.31	.000
Förderbedarf Mathematik[2)]	**–4.20**	0.58	.000
Klassenebene			
IQ-Mittelwert der Klassen	**0.13**	0.05	.014
Anteil Knaben in den Klassen (%)	**0.09**	0.03	.001
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (%)	**–0.04**	0.02	.021
	Var	***SD***	***p***
Zufällige Effekte			
σ^2_e (Fehlervarianz L1)	22.03	4.69	
σ^2_{u0} (Fehlervarianz L2)	2.81	1.68	.000
Erklärte Varianz L1 (%)[3)]		66.05	
Erklärte Varianz L2 (%)[3)]		16.71	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Stichprobe^{IQstreng}: Follow-up 2

Beim Follow-up 2 beträgt die durchschnittliche Mathematikleistung für die Stichprobe^{IQstreng} 44.92 Punkte (vgl. Modell 6 in Tabelle 40) und ist somit um 0.42 Punkte höher als in der Stichprobe^{alle}. Auch der Regressionskoeffizient für das Vorwissen ist mit einem Wert von 0.96 leicht höher (SE = 0.06, p = .000), dieser Zusammenhang ist für diese Untersuchungsstichprobe jedoch in allen Klassen gleichförmig, d.h., es müssen keine Random Slopes berücksichtigt werden. Der standardisierte Regressionskoeffizient für das Vorwissen ist mit 0.58 vergleichbar mit den bisherigen Werten für die Stichprobe^{IQstreng}. Der Regressionskoeffizient für den IQ ist etwas niedriger (*B* = 0.08, *SE* = 0.02, *p* = .000) verglichen mit dem Modell für den Follow-up 2 der Stichprobe^{alle}. Der standardisierte Regressionskoeffizient für den IQ liegt mit 0.11 zwischen den Werten für den Post-Test und den Follow-up 1. Ebenfalls leicht höher ist der Regressionskoeffizient für das Geschlecht (*B* = 0.96, *SE* = 0.44). Zudem wirkt sich ein Förderbedarf in Mathematik mit einem Regressionskoeffizienten von –4.05 (*SE* = 0.99, *p* = .000) in der Stichprobe^{IQstreng} stärker negativ auf die Mathematikleistung aus (Differenz = 0.48 Punkte).

Tabelle 40: Mehrebenenmodell für die Mathematikleistung beim Follow-up 2 mit Individualprädiktoren und Cross-Level-Interaktion für die Stichprobe^{IQstreng}

	Modell 6		
	B	***SE***	***p***
Intercept	**44.92**	0.47	.000
Individualebene			
Vorwissen Mathematik	**0.96**	0.06	.000
IQ	**0.08**	0.02	.000
Geschlecht[1)]	**0.96**	0.44	.028
Förderbedarf Mathematik[2)]	**–4.05**	0.99	.000
Einschränkungen im Deutschverstehen[3)]	**–1.21**	0.56	.029
Cross-Level			
Förderbedarf Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen	**–0.10**	0.04	.017
	Var	***SD***	***p***
Zufällige Effekte			
σ^2_e (Fehlervarianz L1)	36.65	6.05	
σ^2_{u0} (Fehlervarianz L2)	4.54	2.13	.000
σ^2_{u4} (Förderbedarf Mathematik)	9.33	**3.05**	.013
Erklärte Varianz L1 (%)[4)]		58.67[5)]	
Erklärte Varianz L2 (%)[4)]		11.34[5)]	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; [5)] Varianzaufklärung basiert auf Modell ohne Random Slope: Varianzkomponente L1 = 37.89, Varianzkomponente L2 = 6.02; *Var* = Varianzkomponente

Allerdings zeigt sich auch in der Stichprobe^{IQstreng}, dass dieser Zusammenhang nicht in allen Klassen gleich ausfällt (*SD* = 3.05, *p* = .013). Demnach weisen 95% aller Klassen einen Regressionskoeffizienten für den Förderbedarf in Mathematik auf, der zwischen –10.03 und 1.93 liegt. Auch hier kann dieser Unterschied zwischen den Klassen teilweise durch den Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen erklärt werden (B = –0.10, SE = 0.04, p = .017). Demnach führt ein um einen Prozentpunkt höherer Anteil an Kindern mit Einschränkungen im Deutschverstehen im Vergleich zum Durchschnittswert für die Stichprobe dazu, dass sich der Zusammenhang zwischen Förderbedarf in Mathematik und Mathematikleistung beim Follow-up 2 ohne Berücksichtigung der Random Slope um –0.10 Punkte verstärkt (–4.05 – 0.10 = –4.15). Die Abbildung 11 zeigt grafisch den unterschiedlichen Einfluss des Anteils Kinder mit Einschränkungen im Deutschverstehen in den Klassen auf den Zusammenhang zwischen dem Förderbedarf in Mathematik und der Mathematikleistung beim Follow-up 2 in der Stichprobe^{IQstreng}. Zudem haben sich wie schon beim Post-Test auch beim Follow-up 2 Einschränkungen im Deutschverstehen als signifikanter Individualprädiktor in der Stichprobe^{IQstreng} erwiesen (B = –1.21, SE = 0.56, p = .029). Das Modell 6 erklärt auf Individualebene 58.67% der Varianz und auf Klassenebene 11.34% der Varianz. Die Varianzaufklärung beim Follow-up 2 ist in der Stichprobe^{IQstreng} somit auf beiden Ebenen grösser im Vergleich zur Stichprobe^{alle}.

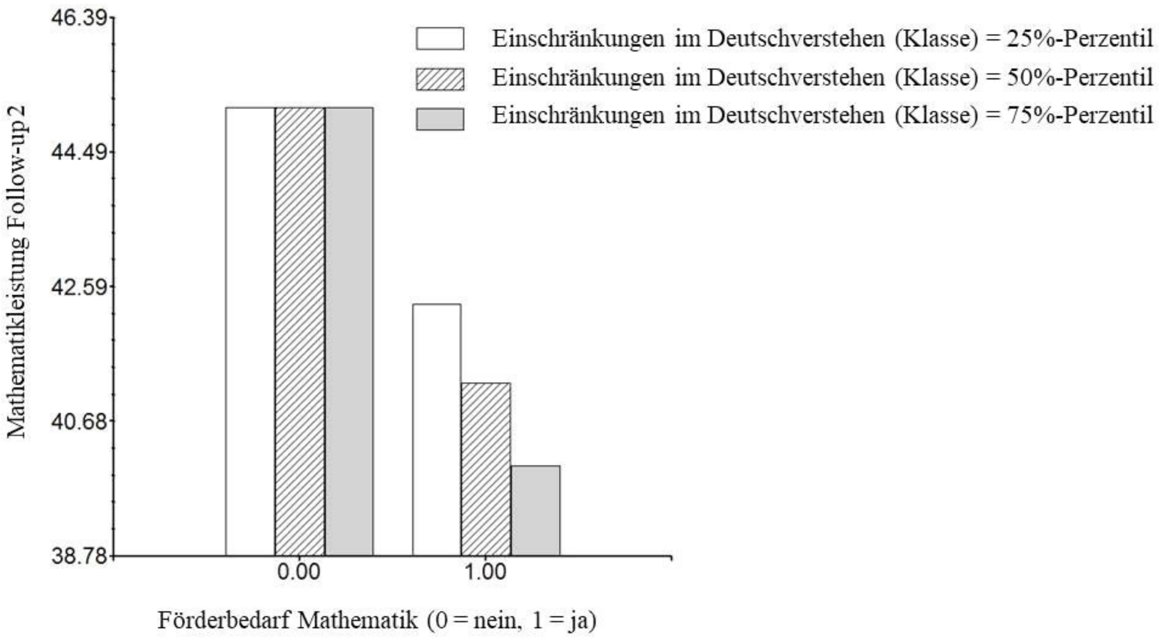


Abbildung 11: Zusammenhang zwischen Mathematikleistung beim Follow-up 2 und Förderbedarf in Mathematik in Abhängigkeit des Anteils Kinder mit Einschränkungen im Deutschverstehen in den Klassen für die Stichprobe^{IQstreng}

Für die Stichprobe^{IQstreng} ergeben sich in den Modellen einige Parallelen im Vergleich zur Stichprobe^{alle}, sodass an dieser Stelle die abweichenden Ergebnisse zusammengefasst werden. So wirken sich in der Stichprobe^{IQstreng} sowohl beim Post-Test als auch beim Follow-up 2, nicht aber beim Follow-up 1 Einschränkungen im Deutschverstehen auf Individualebene negativ auf die jeweilige Mathematikleistung aus. Zudem zeigt sich beim Post-Test, dass auf Klassenebene auch das Vorwissen der Klasse ein signifikanter Prädiktor für die Mathematikleistung ist, wenn nicht gleichzeitig auch der IQ-Mittelwert der Klassen mitmodelliert wird. Daraus kann geschlossen werden, dass das Vorwissen der Klasse und das Intelligenzniveau der Klasse mit Bezug auf die individuellen Mathematikleistungen einen gemeinsamen Anteil an Varianz erklären können. Zudem zeigt sich für den Follow-up 2, dass nur der Zusammenhang zwischen Förderbedarf in Mathematik und der Mathematikleistung zwischen den Klassen unterschiedlich stark ausfällt. Für das Vorwissen in Mathematik müssen in der Stichprobe^{IQstreng} hingegen keine Random Slopes berücksichtigt werden.

6.3 Wirksamkeit der unterrichtsintegrierten Mathematikförderung und Unterschiede zwischen den Interventionsgruppen

In diesem Kapitel wird erstens untersucht, ob Schülerinnen und Schüler der Gruppe^{Begl} und der Gruppe^{Mat}, die eine unterrichtsintegrierte Mathematikförderung erhalten haben, bezüglich ihrer Mathematikleistung stärker profitieren konnten als Schülerinnen und Schüler, deren Lehrpersonen keine Interventionsmassnahmen eingesetzt haben. Zweitens wird analysiert, ob die Form, mit der die Lehrpersonen angeleitet wurden (intensivere Begleitung vs. schriftliche Materialien und Anleitungen), einen Einfluss auf die Leistungsfortschritte der Schülerinnen und Schüler hat. Die Hypothese hierzu lautet, dass Schülerinnen und Schüler der Gruppe^{Begl}, deren Lehrpersonen eine intensivere Begleitung erhalten haben, ihre Mathematikleistung stärker steigern können als die Schülerinnen und Schüler der Gruppe^{Mat}, deren Lehrpersonen nur die schriftlichen Materialien und Anleitungen zur Verfügung hatten.

Um diese Fragestellungen beantworten zu können, werden in die Mehrebenenmodelle, die im Kapitel 6.2 präsentiert wurden, auf Klassenebene zusätzlich Dummy-Variablen eingefügt, mit denen die Unterschiede zwischen der Gruppe^{Begl}, der Gruppe^{Mat} und der Gruppe^{Kontr} untersucht werden können. Damit die im Kapitel 5.5 festgestellten Pre-Test-Unterschiede zwischen der Gruppe^{Begl}, der Gruppe^{Mat} und der Gruppe^{Kontr} kontrolliert werden können, werden zusätzlich auch noch die Variablen SES, Alter und Einschränkungen im Deutschverstehen auf Individualebene eingefügt, wenn diese nicht schon in den bisherigen Modellen enthalten sind. Es werden zunächst für die Stichprobe^{alle} für die drei Messzeitpunkte (Post-Test, Follow-up 1 und Follow-up 2) jeweils zwei Modelle aufgezeigt. Die beiden Modelle unterscheiden sich nur dahingehend, dass im Modell a jeweils die Gruppe^{Kontr} und im Modell b die Gruppe^{Mat} die Referenzgruppe ist. Somit unterscheiden sich in den Modellen a und b jeweils nur der Intercept und die Regressionskoeffizienten der Dummy-Variablen, die für die Stichprobengruppenzugehörigkeit in die Modelle eingefügt wurden. Der Intercept steht jeweils für den Mittelwert für diejenige

Stichprobengruppe, die im jeweiligen Modell als Referenzgruppe definiert wurde. Die restlichen Werte sind in den Modellen a und b identisch, werden vollständigkeitshalber aber dennoch komplett in den Tabellen abgebildet.

In einem zweiten Schritt werden die entsprechenden Modelle für die Stichprobe$^{\text{IQstreng}}$ dargestellt, wobei insbesondere auf Abweichungen im Vergleich zur Stichprobe$^{\text{alle}}$ eingegangen wird. Für den Post-Test wurde für die Stichprobe$^{\text{IQstreng}}$ für die folgenden Analysen das Modell 4b mit dem Prädiktor „IQ-Mittelwert der Klasse" als Ausgangsmodell gewählt, weil im Modell 4b der gleiche Prädiktor auf Klassenebene berücksichtigt wird wie im Modell 1 beim Post-Test für die Stichprobe$^{\text{alle}}$. Die Beschreibung der Ergebnisse fokussiert im Folgenden auf die Interventionseffekte sowie auf die damit verbundenen Effektstärken und die Varianzaufklärung. Zudem werden die Werte für Prädiktoren thematisiert, die sich deutlich von den Werten unterscheiden, die in den Modellen im Kapitel 6.2 dokumentiert wurden.

6.3.1 Stichprobe$^{\text{alle}}$

Post-Test: Stichprobe^{alle}

Beim Post-Test im dritten Quartal der dritten Klasse zeigt sich für die Stichprobe$^{\text{alle}}$, dass sich der Besuch einer Klasse der Gruppe$^{\text{Mat}}$ im Vergleich zum Besuch einer Klasse der Gruppe$^{\text{Kontr}}$ positiv auf die Mathematikleistung ausgewirkt hat. Demnach weisen Kinder aus der Gruppe$^{\text{Mat}}$, die an der halbjährigen, unterrichtsintegrierten Mathematikförderung teilgenommen haben, unter Berücksichtigung der aufgelisteten Prädiktoren direkt nach der Intervention eine um 1.51 Punkte (SE = 0.69) höhere Mathematikleistung auf als Kinder der Gruppe$^{\text{Kontr}}$ bei einem Signifikanzwert von p = .032 (vgl. Modell 7a in Tabelle 41). Die Effektstärke für diesen Unterschied beträgt 0.29, was gemäss Cohen (1988) einem kleinen Effekt entspricht. Da in der vorliegenden Arbeit eine Intervention untersucht wird, die im natürlichen Kontext der Schule umgesetzt wurde, wird im Rahmen der Diskussion noch zu erörtern sein, inwiefern es sich hier dennoch um einen bedeutsamen Effekt handelt (Trautwein et al., 2009; vgl. Kap. 7.1.1). Für die Gruppe$^{\text{Begl}}$ hingegen ist der Vergleich mit der Gruppe$^{\text{Kontr}}$ nicht signifikant (B = –0.69, SE = 0.68, p = .315), was bedeutet, dass die Intervention in der Gruppe$^{\text{Begl}}$ nicht wirksam war. Dem Modell 7b kann zusätzlich entnommen werden, dass die Gruppe$^{\text{Begl}}$ bei einem Regressionskoeffizienten von –2.19 beim Post-Test signifikant schlechter abgeschnitten hat als die Gruppe$^{\text{Mat}}$ (SE = 0.77, p = .006). Damit kann die Hypothese der ersten Fragestellung beim Post-Test für die Stichprobe$^{\text{alle}}$ nur für die Gruppe$^{\text{Mat}}$, nicht aber für die Gruppe$^{\text{Begl}}$ bestätigt werden. Zudem muss die Hypothese für die zweite Fragestellung mit Blick auf die Wirksamkeit der intensiver begleiteten Intervention beim Post-Test für die Stichprobe$^{\text{alle}}$ abgelehnt werden. Die Modelle 7a und 7b erklären auf Individualebene 63.07% und auf Klassenebene 16.19% der Varianz. Im Vergleich zum Modell ohne die Dummy-Variablen für die Stichprobengruppenzugehörigkeit (nicht dargestellt) können die Modelle 7a und 7b auf Individualebene zusätzliche 1.99% der Varianz erklären, auf Klassenebene beträgt dieser Wert 2.10%. Zudem sind die Modelle 7a und 7b signifikant besser als das Modell ohne die Prädiktoren für die Stichprobengruppenzugehörigkeit ($\chi^2(2, N_{L1} = 851, N_{L2} = 58) = 8.57, p = .014$).

Tabelle 41: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Post-Test für die Stichprobe^{alle} in den Modellen 7a (Referenzgruppe: Gruppe^{Kontr}) und 7b (Referenzgruppe: Gruppe^{Mat})

	Modell 7a			Modell 7b		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**25.64**	0.52	.000	**27.15**	0.59	.000
Individualebene						
Vorwissen Mathematik	**0.80**	0.04	.000	**0.80**	0.04	.000
IQ	**0.08**	0.01	.000	**0.08**	0.01	.000
SES	0.08	0.17	.627	0.08	0.17	.627
Alter (in Monaten)	0.02	0.03	.587	0.02	0.03	.587
Geschlecht[1)]	**1.31**	0.39	.001	**1.31**	0.39	.001
Förderbedarf Mathematik[2)]	**–4.09**	0.62	.000	**–4.09**	0.62	.000
Einschränkungen im Deutschverstehen[3)]	–0.67	0.43	.119	–0.67	0.43	.119
Klassenebene						
IQ-Mittelwert der Klassen	**0.15**	0.05	.007	**0.15**	0.05	.007
Gruppe^{Begl} [4)]	–0.69	0.68	.315			
Gruppe^{Mat} [4)]	**1.51**	0.69	.032			
Gruppe^{Begl} [5)]				**–2.19**	0.77	.006
Gruppe^{Kontr} [5)]				**–1.51**	0.69	.032
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	26.66	5.16		26.66	5.16	
σ^2_{u0} (Fehlervarianz L2)	3.30	1.82	.000	3.30	1.82	.000
Erklärte Varianz L1 (%)[6)]		63.07			63.07	
Erklärte Varianz L2 (%)[6)]		16.19			16.19	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Im Gegensatz zum Modell 1 (vgl. Tabelle 37), in dem Prädiktoren mit einem signifikanten Einfluss auf die Mathematikleistung identifiziert wurden, zeigt sich in den Modellen 7a und 7b unter zusätzlicher Berücksichtigung der Kontrollvariablen SES, Alter und Einschränkungen im

Deutschverstehen sowie der Dummy-Variablen für die Stichprobengruppenzugehörigkeit, dass die genannten Kontrollvariablen keinen signifikanten Einfluss auf die Mathematikleistung beim Post-Test haben (SES: B = 0.17, SE = 0.17, p = .325; Alter: B = 0.05, SE = 0.04, p = .196; Einschränkungen im Deutschverstehen: B = –0.76, SE = 0.43, p = .08; vgl. Tabelle 41). Der Regressionskoeffizient für den Förderbedarf in Mathematik ist im Vergleich zum Modell 1 angestiegen (vgl. Tabelle 37), für die restlichen Prädiktoren ergeben sich keine Unterschiede. Die Modelle 7a und 7b erklären auf beiden Ebenen etwas mehr Varianz als das Modell 1.

Follow-up 1: Stichprobe^{alle}

Für die Stichprobe$^{\text{alle}}$ zeigt sich auch für den Follow-up 1 am Ende der dritten Klasse, dass die Gruppe$^{\text{Mat}}$ unter Berücksichtigung der in den Modellen 8a und 8b eingefügten Prädiktoren immer noch signifikant höhere Mathematikleistungen aufweist als die Gruppe$^{\text{Kontr}}$ (B = 1.71, SE = 0.72, p = .020; vgl. Modell 8a in Tabelle 42) und die Schülerinnen und Schüler dieser Interventionsgruppe demnach auch ca. drei Monate nach der Intervention von der mathematischen Förderung profitiert haben. Dabei ergibt sich im Vergleich zum Post-Test eine höhere Effektstärke von 0.40, die gemäss Cohen (1988) aber immer noch als klein eingeschätzt werden muss. Für die Gruppe$^{\text{Begl}}$ zeigt sich auch für den Follow-up 1, dass diese nicht signifikant besser abschneidet als die Gruppe$^{\text{Kontr}}$ (B = 0.37, SE = 0.56, p = .508) und somit kein Interventionseffekt vorliegt. Nicht signifikant ist beim Follow-up 1 im Vergleich zum Post-Test hingegen der Vergleich zwischen der Gruppe$^{\text{Begl}}$ und der Gruppe$^{\text{Mat}}$ (B = –1.34, SE = 0.79, p = .096; vgl. Modell 8b in Tabelle 42). Demnach kann die Hypothese der ersten Fragestellung für die Stichprobe$^{\text{alle}}$ auch beim Follow-up 1 nur für die Gruppe$^{\text{Mat}}$ bestätigt werden, die Hypothese der zweiten Fragestellung muss hingegen auch hier verworfen werden. Das Hinzufügen der Dummy-Variablen für die Stichprobengruppenzugehörigkeit hat zu einer signifikanten Modellverbesserung beim Follow-up 1 geführt ($\chi^2(2, N_{\text{L1}} = 856, N_{\text{L2}} = 58) = 6.16, p = .045$), wobei im Vergleich zum Modell ohne Dummy-Variablen (nicht dargestellt) auf Individualebene 1.04% und auf Klassenebene 1.01% zusätzliche Varianz aufgeklärt werden kann. Die Varianzaufklärung fällt auf Individualebene mit 66.26% höher, auf Klassenebene mit 14.74% hingegen niedriger aus als beim Post-Test für diese Untersuchungsstichprobe.

Tabelle 42: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Follow-up 1 für die Stichprobe^{alle} in den Modellen 8a (Referenzgruppe: Gruppe^{Kontr}) und 8b (Referenzgruppe: Gruppe^{Mat})

	Modell 8a			**Modell 8b**		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**27.30**	0.43	.000	**29.01**	0.66	.000
Individualebene						
Vorwissen Mathematik	**0.81**	0.05	.000	**0.81**	0.05	.000
IQ	**0.05**	0.02	.002	**0.05**	0.02	.002
SES	–0.01	0.14	.923	–0.01	0.14	.923
Alter (in Monaten)	**–0.09**	0.03	.008	**–0.09**	0.03	.008
Geschlecht[1)]	**1.52**	0.31	.000	**1.52**	0.31	.000
Förderbedarf Mathematik[2)]	**–4.25**	0.58	.000	**–4.25**	0.58	.000
Einschränkungen im Deutschverstehen[3)]	–0.46	0.39	.248	–0.46	0.39	.248
Klassenebene						
IQ-Mittelwert der Klassen	**0.14**	0.05	.005	**0.14**	0.05	.005
Anteil Knaben in den Klassen (%)	**0.07**	0.03	.015	**0.07**	0.03	.015
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (%)	–0.03	0.02	.132	–0.03	0.02	.132
Gruppe^{Begl} [4)]	0.37	0.56	.508			
Gruppe^{Mat} [4)]	**1.71**	0.72	.020			
Gruppe^{Begl} [5)]				–1.34	0.79	.096
Gruppe^{Kontr} [5)]				**–1.71**	0.72	.020
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	22.65	4.76		22.65	4.76	
σ^2_{u0} (Fehlervarianz L2)	2.66	1.63	.000	2.66	1.63	.000
Erklärte Varianz L1 (%)[6)]		66.26			66.26	
Erklärte Varianz L2 (%)[6)]		14.74			14.74	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Minimale Unterschiede im Vergleich zum Modell 2 für die Stichprobe^{alle} (vgl. Tabelle 37) ergeben sich in den Modellen 8a und 8b mit Bezug auf die Höhe der Regressionskoeffizienten des Geschlechts (B = 1.52, SE = 0.31, p = .000), des Förderbedarfs in Mathematik (B = –4.25, SE = 0.58, p = .000), des IQ-Mittelwerts der Klassen (B = 0.14, SE = 0.05, p = .005) und des Anteils Knaben in den Klassen (B = 0.07, SE = 0.03, p = .015). Für den IQ-Mittelwert der Klassen zeigt sich zudem, dass die Fehlerwahrscheinlichkeit geringer, beim Anteil Knaben in den Klassen hingegen höher ausfällt als im Modell 2. Der grösste Unterschied bezieht sich jedoch darauf, dass unter zusätzlicher Berücksichtigung der Dummy-Variablen für die Stichprobenzugehörigkeit, des SES und insbesondere der Einschränkungen im Deutschverstehen als Individualprädiktor der Prädiktor „Anteil Kinder mit Einschränkungen im Deutschverstehen" auf Klassenebene nicht mehr signifikant ist (B = –0.03, SE = 0.02, p = .132). Die zusätzlich hinzugefügten Kontrollvariablen auf Individualebene haben beim Follow-up 1 keinen signifikanten Einfluss auf die Mathematikleistung (SES: B = –0.01, SE = 0.14, p = .923; Einschränkungen im Deutschverstehen: B = –0.46, SE = 0.39, p = .248). Die Varianzaufklärung ist in den Modellen 8a und 8b zudem auf beiden Ebenen etwas höher als im Modell 2.

Follow-up 2: Stichprobe^{alle}

Auch beim Follow-up 2 zeigt sich für die Stichprobe$^{\text{alle}}$, dass die Schülerinnen und Schüler der Gruppe$^{\text{Mat}}$ am Ende der vierten Klasse über signifikant höhere Mathematikleistungen verfügen als die Schülerinnen und Schüler der Gruppe$^{\text{Kontr}}$ (B = 1.92, SE = 0.69, p = .007; vgl. Modell 9a in Tabelle 43) und somit auch über ein Jahr nach der unterrichtsintegrierten Mathematikförderung von dieser profitiert zu haben scheinen. Die Effektstärke für diesen Unterschied liegt beim Follow-up 2 bei 0.32, es handelt sich somit auch hier um einen als klein einzuschätzenden Effekt (Cohen, 1988). Für die Gruppe$^{\text{Begl}}$ lässt sich hingegen auch beim Follow-up 2 kein Effekt für die mathematische Förderung feststellen (B = –0.35, SE = 0.89, p = .694). Zudem schneidet die Gruppe$^{\text{Begl}}$ wie auch beim Post-Test signifikant schlechter ab als die Gruppe$^{\text{Mat}}$ (B = –2.28, SE = 0.88, p = .012; vgl. Modell 9b in Tabelle 43). Auch für die Stichprobe$^{\text{alle}}$ muss daher die Hypothese der zweiten Fragestellung für den Follow-up 2 verworfen werden, und die Hypothese der ersten Fragestellung kann nur für die Gruppe$^{\text{Mat}}$ bestätigt werden. Das Hinzufügen der Dummy-Variablen für die Stichprobengruppenzugehörigkeit hat zu einer signifikanten Modellverbesserung beim Follow-up 2 geführt ($\chi^2(2, N_{L1} = 818, N_{L2} = 58) = 6.87$, p = .031), wobei im Vergleich zum Modell ohne Dummy-Variablen (nicht dargestellt) auf Individualebene 1.27% und auf Klassenebene 1.37% zusätzliche Varianz aufgeklärt werden kann. Insgesamt erklären die Modelle 9a und 9b auf Individualebene 58.41% und auf Klassenebene 11.20% der Varianz. Beide Werte sind somit niedriger als beim Post-Test und beim Follow-up 1.

Im Vergleich zum Modell 3 (vgl. Tabelle 37) ergeben sich für die Modelle 9a und 9b nur minimale Abweichungen, sodass diese nicht weiter ausgeführt werden. Die zusätzlich hinzugefügten Kontrollvariablen auf Individualebene sind auch beim Follow-up 2 keine signifikanten Prädiktoren für die Mathematikleistung (SES: B = 0.04, SE = 0.23, p = .861; Alter: B = –0.07, SE = 0.04, p = .117; Einschränkungen im Deutschverstehen B = –0.57, SE = 0.56, p = .306). Die

Varianzaufklärung ist in den Modellen 9a und 9b auf beiden Ebenen zudem leicht höher als im Modell 3.

Tabelle 43: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Follow-up 2 für die Stichprobe^{alle} in den Modellen 9a (Referenzgruppe: Gruppe^{Kontr}) und 9b (Referenzgruppe: Gruppe^{Mat})

	Modell 9a			**Modell 9b**		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**44.03**	0.62	.000	**45.96**	0.60	.000
Individualebene						
Vorwissen Mathematik	**0.94**	0.07	.000	**0.94**	0.07	.000
IQ	**0.09**	0.02	.000	**0.09**	0.02	.000
SES	0.04	0.23	.861	0.04	0.23	.861
Alter (in Monaten)	–0.07	0.04	.117	–0.07	0.04	.117
Geschlecht[1)]	**0.91**	0.43	.033	**0.91**	0.43	.033
Förderbedarf Mathematik[2)]	**–3.59**	0.88	.000	**–3.59**	0.88	.000
Einschränkungen im Deutschverstehen[3)]	–0.57	0.56	.306	–0.57	0.56	.306
Cross-Level-Interaktion						
Förderbedarf Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen	**–0.10**	0.04	.006	**–0.10**	0.04	.006
Klassenebene						
Gruppe^{Begl} [4)]	–0.35	0.89	.694			
Gruppe^{Mat} [4)]	**1.92**	0.69	.007			
Gruppe^{Begl} [5)]				**–2.28**	0.88	.012
Gruppe^{Kontr} [5)]				**–1.92**	0.69	.007
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	36.41	6.03		36.41	6.03	
σ^2_{u0} (Fehlervarianz L2)	6.01	2.45	.000	6.01	2.45	.000
σ^2_{u1} (Vorwissen Mathematik)	0.07	**0.27**	.003	0.07	**0.27**	.003
σ^2_{u6} (Förderbedarf Mathematik)	17.43	**4.18**	.001	17.43	**4.18**	.001
Erklärte Varianz L1 (%)[6)]		58.41[7)]			58.41[7)]	
Erklärte Varianz L2 (%)[6)]		11.20[7)]			11.20[7)]	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutsch verstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; [7)] Varianzaufklärung basiert auf Modell ohne Random Slopes: Varianzkomponente L1 = 39.13, Varianzkomponente L2 = 5.69; *Var* = Varianzkomponente

Zusammenfassend kann für die Stichprobe^{alle} daher festgestellt werden, dass die Gruppe^{Mat} von der unterrichtsintegrierten Mathematikförderung profitiert hat und sich dieser Effekt ca. drei Monate nach der Intervention beim Follow-up 1 und auch fast eineinhalb Jahre nach der Intervention beim Follow-up 2 noch zeigt, sodass von einer langfristigen Wirkung der Förderung für diese Interventionsgruppe ausgegangen werden kann. Die diesbezüglichen Effektstärken von 0.29 (Post-Test), 0.40 (Follow-up 1) und 0.32 (Follow-up 2) müssen gemäss Cohen (1988) jedoch als eher klein eingeschätzt werden. Inwiefern es sich hier dennoch um bedeutsame Effekte handelt, wird im Rahmen der Diskussion mit Verweis auf Effektstärken aus vergleichbaren Interventionsstudien noch besprochen (vgl. Kap. 7.1.1). Die Hypothese der ersten Fragestellung kann somit für die Gruppe^{Mat} bestätigt werden. Für die Gruppe^{Begl}, in der die Lehrpersonen im Vergleich zur Gruppe^{Mat} eine intensivere Begleitung erhalten haben, zeigt sich hingegen bei allen Messzeitpunkten, dass diese im Vergleich zur Gruppe^{Kontr} nicht über bessere Mathematikleistungen verfügt und somit die mathematische Förderung in der Gruppe^{Begl} nicht wirksam war. Daher kann die Hypothese der ersten Fragestellung für die Gruppe^{Begl} nicht bestätigt werden. Zudem zeigt sich beim Post-Test und beim Follow-up 2, dass die Gruppe^{Begl} signifikant schlechter abschneidet als die Gruppe^{Mat}. Somit muss auch die zweite Hypothese klar verworfen werden. Aufgrund der bisherigen Analysen lässt sich nicht sagen, wieso sich diese unterschiedlichen Effekte für die unterrichtsintegrierte Mathematikförderung in den Interventionsgruppen Gruppe^{Begl} und Gruppe^{Mat} zeigen.

6.3.2 Stichprobe^{IQstreng}

Post-Test: Stichprobe^{IQstreng}

Beim Post-Test direkt im Anschluss an die mathematische Intervention zeigt sich auch für die Stichprobe^{IQstreng}, dass die Gruppe^{Mat} im Vergleich zur Gruppe^{Kontr} eine signifikant höhere Mathematikleistung aufweist und somit von der unterrichtsintegrierten Mathematikförderung profitiert hat (B = 1.50, SE = 0.71, p = .039; vgl. Modell 10a in Tabelle 44). Die Effektstärke für diesen Unterschied beträgt in der Stichprobe^{IQstreng} beim Post-Test ebenfalls 0.29, was gemäss Cohen (1988) einem kleinen Effekt entspricht. Für die Gruppe^{Begl} hingegen lässt sich auch in der Stichprobe^{IQstreng} kein signifikanter Unterschied im Vergleich zur Gruppe^{Kontr} feststellen (B = –0.70, SE = 0.69, p = .313). Zudem schneidet auch in dieser Untersuchungsstichprobe die Gruppe^{Begl} signifikant schlechter ab als die Gruppe^{Mat} (B = –2.20, SE = 0.77, p = .006; vgl. Modell 10b in Tabelle 44). Damit kann die Hypothese der ersten Fragestellung beim Post-Test auch für die Stichprobe^{IQstreng} nur für die Gruppe^{Mat}, nicht aber für die Gruppe^{Begl} bestätigt werden. Zudem muss die Hypothese für die zweite Fragestellung zur Wirksamkeit der intensiveren Begleitung für den Post-Test auch für die Stichprobe^{IQstreng} abgelehnt werden.

Tabelle 44: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Post-Test für die Stichprobe^{IQstreng} in den Modellen 10a (Referenzgruppe: Gruppe^{Kontr}) und 10b (Referenzgruppe: Gruppe^{Mat})

	Modell 10a			Modell 10b		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**25.86**	0.52	.000	**27.35**	0.59	.000
Individualebene						
Vorwissen Mathematik	**0.80**	0.05	.000	**0.80**	0.05	.000
IQ	**0.08**	0.02	.000	**0.08**	0.02	.000
SES	0.17	0.17	.325	0.17	0.17	.325
Alter (in Monaten)	0.05	0.04	.196	0.05	0.04	.196
Geschlecht[1)]	**1.29**	0.39	.001	**1.29**	0.39	.001
Förderbedarf Mathematik[2)]	**–3.91**	0.59	.000	**–3.91**	0.59	.000
Einschränkungen im Deutschverstehen[3)]	–0.76	0.43	.080	–0.76	0.43	.080
Klassenebene						
IQ-Mittelwert der Klassen	**0.14**	0.05	.010	**0.14**	0.05	.010
Gruppe^{Begl} [4)]	–0.70	0.69	.313			
Gruppe^{Mat} [4)]	**1.50**	0.71	.039			
Gruppe^{Begl} [5)]				**–2.20**	0.77	.006
Gruppe^{Kontr} [5)]				**–1.50**	0.05	.039
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	25.83	5.08		25.83	5.08	
σ^2_{u0} (Fehlervarianz L2)	3.39	1.84	.000	3.39	1.84	.000
Erklärte Varianz L1 (%)[6)]		63.17			63.17	
Erklärte Varianz L2 (%)[6)]		17.95			17.95	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Der Intercept ist mit 25.86 Punkten beim Post-Test in der Stichprobe^{IQstreng} etwas höher als in der Stichprobe^{alle} (vgl. Tabelle 41). Dies kann auch hier durch den Ausschluss von Kindern, deren Alter die IQ-Altersnorm übersteigt und die u.a. über weniger mathematisches Vorwissen im Vergleich zur Gesamtstichprobe verfügen, erklärt werden (vgl. Kapitel 5.5). Der Einfluss des Förderbedarfs in Mathematik fällt zudem im Vergleich zur Stichprobe^{alle} etwas weniger stark aus ($B = –3.91$, $SE = 0.59$, $p =.000$). Die Modelle 10a und 10b erklären auf Individualebene 63.17% der Varianz und auf Klassenebene 17.95%, somit sind beide Werte (leicht) höher als in der Stichprobe^{alle}. Im Vergleich zu den Modellen ohne die Prädiktoren für die Stichprobenzugehörigkeit (nicht dargestellt) ergibt dies eine zusätzlich erklärte Varianz von 1.94% auf Individualebene und 2.01% auf Klassenebene, wobei die Modelle 10a und 10b signifikant besser sind als das Modell ohne die Dummy-Variablen für die Stichprobengruppenzugehörigkeit ($\chi^2(2, N_{L1} = 806, N_{L2} = 58) = 8.37, p = .015$).

Im Gegensatz zum Modell 4b der Stichprobe^{IQstreng} (vgl. Tabelle 38), in dem die Prädiktoren der Mathematikleistung identifiziert wurden, zeigt sich, dass in den Modellen 10a und 10b unter zusätzlicher Berücksichtigung des SES, des Alters und der Dummy-Variablen für die Stichprobengruppenzugehörigkeit der Prädiktor „Einschränkungen im Deutschverstehen" auf Individualebene nur einen tendenziellen Einfluss auf die Mathematikleistung beim Post-Test hat ($B = –0.76$, $SE = 0.43$, $p = .080$; vgl. Tabelle 44). Die Kontrollvariablen SES ($B = 0.17$, $SE = 0.17$, $p = .325$) und Alter ($B = 0.05$, $SE = 0.04$, $p = .196$) sind auch in der Stichprobe^{IQstreng} keine signifikanten Prädiktoren für die Mathematikleistung beim Post-Test. Minimale Unterschiede zum Modell 4b ergeben sich für die Höhe der Regressionskoeffizienten des Geschlechts ($B = 1.29$, $SE = 0.39$, $p = .001$), des Förderbedarfs in Mathematik ($B = –3.91$, $SE = 0.59$, $p =.000$) und des IQ-Mittelwerts der Klasse ($B = 0.14$, $SE = 0.05$, $p =.010$). Auch die Varianzaufklärung ist im Vergleich zum Modell 4b auf beiden Ebenen leicht höher.

Follow-up 1: Stichprobe^{IQstreng}

Für die Stichprobe^{IQstreng} zeigt sich für den Follow-up 1 ebenfalls, dass die Gruppe^{Mat} signifikant von der Förderung profitieren konnte ($B = 1.54$, $SE = 0.73$, $p = .038$; vgl. Modell 11a in Tabelle 45), nicht jedoch die Gruppe^{Begl} ($B = 0.37$, $SE = 0.57$, $p = .517$). In der Stichprobe^{IQstreng} ist der Koeffizient für die Gruppe^{Mat} zudem mit 1.54 Punkten kleiner als im Modell 8a für die Stichprobe^{alle} (vgl. Tabelle 42). Entsprechend ist auch die Effektstärke für diesen Unterschied mit 0.33 beim Follow-up 1 für die Stichprobe^{IQstreng} etwas geringer. Weiter ist der Unterschied zwischen der Gruppe^{Mat} und der Gruppe^{Begl} auch in der Stichprobe^{IQstreng} beim Follow-up 1 nicht signifikant ($B = –1.18$, $SE = 0.78$, $p = .138$; vgl. Modell 11b in Tabelle 45). Demnach kann die Hypothese der ersten Fragestellung auch für die Stichprobe^{IQstreng} nur für die Gruppe-^{Mat} bestätigt werden, die Hypothese der zweiten Fragestellung muss auch hier verworfen werden. Die Varianzaufklärung in den Modellen 11a und 11b beträgt auf Individualebene 66.29% und auf Klassenebene 16.96%. Beide Werte sind leicht höher als in der Stichprobe^{alle} beim Follow-up 1. Die Modelle 11a und 11b sind nur tendenziell besser als das Modell ohne die Dummy-Variablen für die Stichprobengruppenzugehörigkeit ($\chi^2(2, N_{L1} = 811, N_{L2} = 58) = 5.00$, $p = .080$), sodass durch das Hinzufügen der Dummy-Variablen auf Individualebene nur 0.75% und auf Klassenebene nur 0.72% zusätzlicher Varianz aufgeklärt wird.

Tabelle 45: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Follow-up 1 für die Stichprobe^{IQstreng} in den Modellen 11a (Referenzgruppe: Gruppe^{Kontr}) und 11b (Referenzgruppe: Gruppe^{Mat})

	Modell 11a			**Modell 11b**		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**27.62**	0.44	.000	**29.17**	0.66	.000
Individualebene						
Vorwissen Mathematik	**0.82**	0.04	.000	**0.82**	0.04	.000
IQ	**0.04**	0.01	.004	**0.04**	0.01	.004
SES	0.08	0.13	.554	0.08	0.13	.554
Alter (in Monaten)	**–0.10**	0.04	.017	**–0.10**	0.04	.017
Geschlecht[1)]	**1.44**	0.32	.000	**1.44**	0.32	.000
Förderbedarf Mathematik[2)]	**–4.22**	0.57	.000	**–4.22**	0.57	.000
Einschränkungen im Deutschverstehen[3)]	–0.48	0.40	.238	–0.48	0.40	.238
Klassenebene						
IQ-Mittelwert der Klassen	**0.15**	0.05	.004	**0.15**	0.05	.004
Anteil Knaben in den Klassen (%)	**0.07**	0.03	.014	**0.07**	0.03	.014
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (%)	–0.03	0.02	.087	–0.03	0.02	.087
Gruppe^{Begl} [4)]	0.37	0.57	.517			
Gruppe^{Mat} [4)]	**1.54**	0.73	.038			
Gruppe^{Begl} [5)]				–1.18	0.78	.138
Gruppe^{Kontr} [5)]				**–1.54**	0.73	.038
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	22.02	4.69		22.02	4.69	
σ^2_{u0} (Fehlervarianz L2)	2.64	1.63	.000	2.64	1.63	.000
Erklärte Varianz L1 (%)[6)]		66.29			66.29	
Erklärte Varianz L2 (%)[6)]		16.96			16.96	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

Im Vergleich zum Modell 5 (vgl. Tabelle 39) hat der Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen in den Modellen 11a und 11b nur einen tendenziellen Einfluss auf die Mathematikleistung (B = –0.03, SE = 0.02, p = .086), wenn zusätzlich Einschränkungen im Deutschverstehen als Individualprädiktor (B = –0.48, SE = 0.40, p = .238) sowie der SES (B = 0.08, SE = 0.13, p = .554) und die Dummy-Variablen für die Stichprobengruppenzugehörigkeit mitmodelliert werden. Geringe Unterschiede ergeben sich für die Höhe der Regressionskoeffizienten des Alters (*B* = –0.10, *SE* = 0.04, *p* = .017), des Geschlechts (*B* = 1.44, *SE* = 0.32, *p* =.000), des Förderbedarfs in Mathematik (*B* = –4.22, *SE* = 0.57, *p* =.000), des IQ-Mittelwerts der Klasse (*B* = 0.15, *SE* = 0.05, *p* =.004) und des Anteils Knaben in den Klassen (*B* = 0.07, *SE* = 0.03, *p* =.014), wobei sich teilweise auch das Signifikanzniveau verändert. Zudem ist die Varianzaufklärung im Vergleich zum Modell 5 auf Individual- und auf Klassenebene leicht höher.

Follow-up 2: Stichprobe^{IQstreng}

Auch beim Follow-up 2 am Ende der vierten Klasse zeigt sich für die Stichprobe$^{\text{IQstreng}}$, dass die Schülerinnen und Schüler der Gruppe$^{\text{Mat}}$ über signifikant höhere Mathematikleistungen verfügen im Vergleich zur Gruppe$^{\text{Kontr}}$ (*B* = 1.88, *SE* = 0.72, *p* = .012; vgl. Modell 12a in Tabelle 46). Die Effektstärke liegt hier bei 0.30. Die Gruppe$^{\text{Begl}}$ hingegen unterscheidet sich nach wie vor nicht signifikant von der Gruppe$^{\text{Kontr}}$ (*B* = –0.55, *SE* = 0.92, *p* = .547). Im Gegensatz zum Follow-up 1, aber in Übereinstimmung mit dem Post-Test schneidet die Gruppe$^{\text{Begl}}$ beim Follow-up 2 signifikant schlechter ab als die Gruppe$^{\text{Mat}}$ (*B* = –2.44, *SE* = 0.87, *p* = .007; vgl. Modell 12b in Tabelle 46). Demnach kann auch für die Stichprobe$^{\text{IQstreng}}$ mehr als ein Jahr nach der Intervention die Hypothese der ersten Fragestellung nur für die Gruppe$^{\text{Mat}}$ bestätigt werden, die Hypothese der zweiten Fragestellung muss hingegen abgelehnt werden. Die Modelle 12a und 12b sind signifikant besser als das Modell ohne die Prädiktoren für die Stichprobengruppenzugehörigkeit ($\chi^2(2, N_{L1} = 778, N_{L2} = 58) = 7.59$, *p* = .022). Im Vergleich zum Modell ohne Dummy-Variablen (nicht dargestellt) erklären die Modelle 12a und 12b auf Individualebene 1.18% und auf Klassenebene 1.27% zusätzliche Varianz, sodass diese Werte etwas geringer ausfallen als in der Stichprobe$^{\text{alle}}$. Insgesamt liegt die Varianzaufklärung in den Modellen 12a und 12b auf Individualebene bei 59.08% und auf Klassenebene bei 11.80% der Varianz. Im Vergleich zur Stichprobe$^{\text{alle}}$ ist der Wert auf Klassenebene geringfügig höher, auf Individualebene hingegen etwas niedriger.

Vergleicht man die Modelle 12a und 12b mit dem Modell 6 (vgl. Tabelle 40) haben Einschränkungen im Deutschverstehen auch hier nur einen tendenziellen Einfluss auf die Mathematikleistungen. Im Vergleich zum Post-Test und zum Follow-up 1 ist dieser Unterschied aber nur knapp nicht signifikant (*B* = –1.04, *SE* = 0.56, *p* = .060). Zudem gibt es leichte Unterschiede in der Höhe der Regressionskoeffizienten für das Geschlecht (*B* = 0.91, *SE* = 0.44, *p* = .037), den Förderbedarf in Mathematik (*B* = –4.07, *SE* = 0.99, *p* = .000) und die Cross-Level-Interaktion (*B* = –0.09, *SE* = 0.04, *p* = .024). Beim Geschlecht und bei der Cross-Level-Interaktion zeigt sich ausserdem, dass die Irrtumswahrscheinlichkeit, dass es sich hier um einen signifikanten Prädiktor handelt, in den Modellen 12a und 12b etwas grösser ist als im Modell 6. Auf Individualebene ist die Varianzaufklärung vergleichbar hoch wie im Modell 6, auf Klassenebene hingegen ist der Wert leicht höher.

Tabelle 46: Wirksamkeit der Förderung und Unterschiede zwischen den Interventionsgruppen beim Follow-up 2 für die Stichprobe^{IQstreng} in den Modellen 12a (Referenzgruppe: Gruppe^{Kontr}) und 12b (Referenzgruppe: Gruppe^{Mat})

	Modell 12a			Modell 12b		
	B	***SE***	***p***	***B***	***SE***	***p***
Intercept	**44.56**	0.64	.000	**46.44**	0.53	.000
Individualebene						
Vorwissen Mathematik	**0.96**	0.06	.000	**0.96**	0.06	.000
IQ	**0.08**	0.02	.000	**0.08**	0.02	.000
SES	0.13	0.23	.577	0.13	0.23	.577
Alter (in Monaten)	–0.07	0.05	.170	–0.07	0.05	.170
Geschlecht[1)]	**0.91**	0.44	.037	**0.91**	0.44	.037
Förderbedarf Mathematik[2)]	**–4.07**	0.99	.000	**–4.07**	0.99	.000
Einschränkungen im Deutschverstehen[3)]	–1.04	0.56	.060	–1.04	0.56	.060
Cross-Level-Interaktion						
Förderbedarf Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen	**–0.09**	0.04	.024	**–0.09**	0.04	.024
Klassenebene						
Gruppe^{Begl} [4)]	–0.55	0.92	.547			
Gruppe^{Mat} [4)]	**1.88**	0.72	.012			
Gruppe^{Begl} [5)]				**–2.44**	0.87	.007
Gruppe^{Kontr} [5)]				**–1.88**	0.72	.012
	Var	***SD***	***p***	***Var***	***SD***	***p***
Zufällige Effekte						
σ^2_e (Fehlervarianz L1)	36.58	6.05		36.58	4.69	
σ^2_{u0} (Fehlervarianz L2)	3.86	1.96	.000	3.86	1.63	.000
σ^2_{u6} (Förderbedarf Mathematik)	10.04	**3.17**	.010	10.04	**3.17**	.010
Erklärte Varianz L1 (%)[6)]		59.08[7)]			59.08[7)]	
Erklärte Varianz L2 (%)[6)]		11.80[7)]			11.80[7)]	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Referenzgruppe: Gruppe^{Mat}; [6)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; [7)] Varianzaufklärung basiert auf Modell ohne Random Slopes: Varianzkomponente L1 = 37.91, Varianzkomponente L2 = 5.56; *Var* = Varianzkomponente

Zusammenfassend lassen sich somit für die Stichprobe^{IQstreng} die gleichen Schlüsse bezüglich der Wirksamkeit der unterrichtsintegrierten Mathematikförderung ziehen wie für die Stichprobe^{alle}. Demnach kann die Hypothese der ersten Fragestellung nur für die Gruppe^{Mat} bestätigt werden, muss für die Gruppe^{Begl} jedoch verworfen werden. Die Effektstärken für die Gruppe^{Mat} von 0.29 (Post-Test), 0.33 (Follow-up 1) und 0.30 (Follow-up 2) müssen auch hier gemäss Cohen (1988) als eher klein eingeschätzt werden. Ein Vergleich mit Effektstärken aus ähnlichen Interventionsstudien wird auch hier im Rahmen der Diskussion thematisiert (vgl. Kap. 7.1.1). Zudem wurde die Hypothese der zweiten Fragestellung auch in der Stichprobe^{IQstreng} widerlegt. Auch hier lassen sich aus den bisherigen Analysen keine Gründe ableiten, wieso es zu unterschiedlichen Effekten bei den beiden Interventionsgruppen Gruppe^{Begl} und Gruppe^{Mat} gekommen ist.

6.4 Variablen auf Individual- und Klassenebene mit Einfluss auf die Wirksamkeit der Mathematikförderung

Da auf Basis der bisherigen Modelle nicht geklärt werden konnte, wieso die Förderung in der Gruppe^{Mat} effektiv war, nicht aber in der Gruppe^{Begl}, wird im Folgenden der Frage nachgegangen, ob einzelne Variablen auf Individual- und Klassenebene sich in den einzelnen Stichprobengruppen unterschiedlich auf die Mathematikleistung ausgewirkt haben. Damit sollen Faktoren identifiziert werden, die möglicherweise in einem Zusammenhang mit der unterschiedlichen Wirksamkeit der Förderung in den beiden Interventionsgruppen stehen. Dafür wurden einerseits Cross-Level-Interaktionen zwischen Individualvariablen und Stichprobengruppenzugehörigkeit (Gruppe^{Begl}, Gruppe^{Mat}, Gruppe^{Kontr}) und andererseits Interaktionsterme zwischen Klassenvariablen und Stichprobengruppenzugehörigkeit den bisherigen Modellen hinzugefügt. Im Folgenden werden nur Modelle für die Stichprobe^{alle} präsentiert, da sich nur geringe Unterschiede in den Modellen für die Stichprobe^{alle} und die Stichprobe^{IQstreng} ergeben haben. Es zeigt sich, dass die Variable „Alter" auf Individualebene sowie die Variable „Anteil Kinder mit Einschränkungen im Deutschverstehen" auf Klassenebene sich in den drei Stichprobengruppen bei einzelnen Messzeitpunkten unterschiedlich auf die Mathematikleistung ausgewirkt haben. Entsprechend können diese Variablen möglicherweise in einem Zusammenhang mit der unterschiedlichen Wirksamkeit der Förderung stehen. Alle anderen Cross-Level-Interaktionen und Interaktionsterme auf Klassenebene mit Bezug auf die Stichprobengruppenzugehörigkeit haben sich als nicht signifikant erwiesen und werden daher nicht thematisiert.

6.4.1 Cross-Level-Modell „Alter × Stichprobengruppe" für den Follow-up 1

Beim Follow-up 1 ergab sich sowohl in der Stichprobe^{alle} als auch in der Stichprobe^{IQstreng} ein negativer Zusammenhang zwischen Alter und Mathematikleistung (vgl. Kapitel 6.3). Unter zusätzlicher Berücksichtigung der Stichprobengruppenzugehörigkeit in Form einer Cross-

Level-Interaktion „Alter × Stichprobengruppenzugehörigkeit“ zeigt sich jedoch in beiden Untersuchungsstichproben, dass dieser Zusammenhang nicht in allen drei Stichprobengruppen gleich ausfällt (vgl. Modell 13 in Tabelle 47).

Tabelle 47: Mehrebenenmodell mit Cross-Level-Interaktion „Alter × Stichprobengruppe“ für die Stichprobe^{alle} beim Follow-up 1

	Modell 13		
	B	***SE***	***p***
Intercept	**28.84**	0.66	.000
Individualebene			
Vorwissen Mathematik	**0.83**	0.05	.000
IQ	**0.05**	0.02	.004
SES	–0.04	0.14	.775
Alter (in Monaten)	0.08	0.04	.063
Alter × Gruppe^{Begl} [1)]	**–0.24**	0.06	.000
Alter × Gruppe^{Kontr} [1)]	**–0.25**	0.07	.000
Geschlecht[2)]	**1.54**	0.30	.000
Förderbedarf Mathematik[3)]	**–4.18**	0.57	.000
Einschränkungen im Deutschverstehen[4)]	–0.56	0.40	.159
Klassenebene			
IQ-Mittelwert der Klassen	**0.15**	0.05	.004
Anteil Knaben in den Klassen (%)	**0.07**	0.03	.012
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (%)	–0.03	0.02	.146
Gruppe^{Begl} [4)]	–1.17	0.78	.139
Gruppe^{Kontr} [4)]	**–1.61**	0.71	.027
	Var	***SD***	***p***
Zufällige Effekte			
σ^2_e (Fehlervarianz L1)	22.31	4.72	
σ^2_{u0} (Fehlervarianz L2)	2.71	1.65	.000
Erklärte Varianz L1 (%)[5)]		66.65	
Erklärte Varianz L2 (%)[5)]		14.71	

[1)] Referenzgruppe: Gruppe^{Mat}; [2)] Referenzgruppe: Mädchen; [3)] Referenzgruppe: kein Förderbedarf in Mathematik; [4)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [5)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; *Var* = Varianzkomponente

In diesem Modell, in dem die Gruppe^{Mat} die Referenzgruppe darstellt, wird ersichtlich, dass für die Gruppe^{Mat} ein tendenziell positiver Zusammenhang bei einem Regressionskoeffizienten von 0.08 (SE = 0.04, p = .063) zwischen Alter und Mathematikleistung besteht. Im Vergleich dazu ergibt sich für die Gruppe^{Begl} ein negativer Regressionskoeffizient von –0.24 (SE = 0.06, p = .000), der höchst signifikant ausfällt. Für die Gruppe^{Begl} resultiert für Kinder, deren Alter das durchschnittliche Alter der Stichprobe^{alle} um einen Monat übersteigt, eine um –0.16 Punkte niedrigere Mathematikleistung (0.08 – 0.24 = –0.16). Auch für die Gruppe^{Kontr} zeigt sich im Vergleich zur Gruppe^{Mat} ein negativer Zusammenhang zwischen Alter und Mathematikleistung (B = –0.25, SE = 0.07, p = .000), der höchst signifikant ist. Für ein um einen Monat älteres Kind fällt die Mathematikleistung somit um –0.17 Punkte niedriger aus (0.08 – 0.25 = –0.17).

Daraus kann gefolgert werden, dass in der Gruppe^{Mat}, in der die unterrichtsintegrierte Förderung bei allen drei Messzeitpunkten wirksam war, auch ältere Kinder beim Follow-up 1 von der Förderung profitieren konnten, nicht aber in der Gruppe^{Begl}, in der die Förderung nicht effektiv war. Weshalb ältere Kinder der Gruppe^{Mat} signifikant von der Förderung profitieren konnten, lässt sich allerdings aufgrund des präsentierten Modells nicht sagen. Mögliche Erklärungen dazu werden im Rahmen der Diskussion thematisiert (vgl. Kap. 7.1).

6.4.2 Mehrebenenmodell mit Interaktionsterm „Anteil Kinder mit Einschränkungen im Deutschverstehen × Stichprobengruppe“ für den Follow-up 2

Ebenfalls interessant ist, dass beim Follow-up 2 in beiden Untersuchungsstichproben für die Gruppe^{Begl}, für die die Wirksamkeit der unterrichtsintegrierten Mathematikförderung nicht belegt werden konnte, der Zusammenhang zwischen dem Anteil Kinder mit Einschränkungen im Deutschverstehen und der Mathematikleistung signifikant negativer ausfällt als für die Gruppe^{Kontr} (vgl. Tabelle 48). Zwar hat im Modell 14 der Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen selbst keinen signifikanten Einfluss (B = 0.04, SE = 0.03, p = .130). Der Interaktionsterm „Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen × Gruppe^{Begl}“ zeigt jedoch, dass in der Gruppe^{Begl} sich der Anteil Kinder mit Einschränkungen im Deutschverstehen im Vergleich zur Gruppe^{Kontr} mit einem Regressionskoeffizienten von –1.89 (SE = 0.84, p = .029) signifikant negativer auf die Mathematikleistung beim Follow-up 2 auswirkt. Für die Gruppe^{Mat}, für die sich die unterrichtsintegrierte Förderung als wirksam erwiesen hat, lässt sich hingegen nicht belegen, dass ein hoher Anteil Kinder mit Einschränkungen im Deutschverstehen sich negativ auf die Mathematikleistung auswirkt (B = –0.83, SE = 0.82, p = .315). Mögliche Erklärungen werden im Rahmen der Diskussion besprochen (vgl. Kap. 7.1).

Tabelle 48: Mehrebenenmodell mit Interaktionsterm „Anteil Kinder mit Einschränkungen im Deutschverstehen × Stichprobengruppe" für die Stichprobe^{alle} beim Follow-up 2

	Modell 14		
	B	***SE***	***p***
Intercept	**44.63**	0.51	.000
Individualebene			
Vorwissen Mathematik	**0.95**	0.07	.000
IQ	**0.09**	0.02	.000
Geschlecht[1)]	**0.90**	0.43	.036
Förderbedarf Mathematik[2)]	**–3.60**	0.89	.000
Einschränkungen im Deutschverstehen[3)]	–0.71	0.54	.188
Klassenebene			
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen	0.04	0.03	.130
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen × Gruppe^{Begl} [4)]	**–1.89**	0.84	.029
Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen × Gruppe^{Mat} [4)]	–0.83	0.82	.315
Cross-Level-Interaktion			
Förderbedarf Mathematik × Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen	**–0.10**	0.03	.005
	Var	***SD***	***p***
Zufällige Effekte			
σ^2_e (Fehlervarianz L1)	36.61	6.05	
σ^2_{u0} (Fehlervarianz L2)	6.54	2.56	.000
σ^2_{u1} (Vorwissen Mathematik)	0.08	**0.27**	.003
σ^2_{u4} (Förderbedarf Mathematik)	17.40	**4.17**	.001
Erklärte Varianz L1 (%)[5)]		58.45[6)]	
Erklärte Varianz L2 (%)[5)]		11.29[6)]	

[1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschränkungen im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}; [5)] Varianzaufklärung korrigiert nach Snijders und Bosker (1994) unter Verwendung des harmonischen Mittels für die Berechnung von *n*; [6)] Varianzaufklärung basiert auf Modell ohne Random Slopes: Varianzkomponente L1 = 39.19, Varianzkomponente L2 = 5.58; *Var* = Varianzkomponente

6.5 Wirksamkeit der unterrichtsintegrierten Mathematikförderung bei Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen

Da sich auf Basis der bisher präsentierten Mehrebenenmodelle bezüglich der Wirksamkeit der unterrichtsintegrierten Mathematikförderung ein Bild zeigt, das nicht den Erwartungen entspricht, soll abschliessend noch der Frage nachgegangen werden, ob sich die bisherigen Ergebnisse auch für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen bestätigen lassen. Da pro Klasse nur wenige Schülerinnen und Schüler mit entsprechenden Mathematikleistungen vorhanden sind, können an dieser Stelle jedoch keine Mehrebenenanalysen für die drei Messzeitpunkte berechnet werden. Daher wird im Folgenden mithilfe einer multiplen Regressionsanalyse per Einschlussverfahren ohne Berücksichtigung der hierarchischen Datenstruktur der Frage nachgegangen, inwiefern auch Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen von der unterrichtsintegrierten Förderung im Fach Mathematik profitieren konnten. Es werden die gleichen Individualprädiktoren in der multiplen Regressionsanalyse berücksichtigt wie in den bisher dargestellten Mehrebenenmodellen, ebenso werden Dummy-Variablen für die Gruppenzugehörigkeit eingefügt. Auf Basis der in der Forschung verwendeten Cut-off-Werte für das Vorliegen einer Rechenschwäche (vgl. Kap. 2.1.3) wurden für die Regressionsanalyse alle Schülerinnen und Schüler der Stichprobe$^{\text{alle}}$ berücksichtigt, die beim Pre-Test Mathematikleistungen gezeigt haben, die im untersten Quartil der Stichprobe lagen und somit über unterdurchschnittliche Mathematikleistungen verfügten (≤ 17 Punkte im Pre-Test).

Beim Post-Test zeigen sich für die Gruppe$^{\text{Begl}}$ und die Gruppe$^{\text{Mat}}$ negative Regressionskoeffizienten, wenn nur Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen untersucht werden. Für die Gruppe$^{\text{Begl}}$ beträgt der standardisierte Regressionskoeffizient –.12, es liegt ein tendenziell signifikanter Unterschied zur Gruppe$^{\text{Kontr}}$ vor (p = .065; vgl. Tabelle 49).

Tabelle 49: Multiple Regressionsanalyse für die abhängige Variable Mathematikleistung beim Post-Test für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen beim Pre-Test

Variable	***B***	***SE (B)***	**β**	***t***	***p***
(Konstante)	–5.58	8.66		–0.64	.520
Vorwissen Mathematik	**0.77**	0.12	**.40**	6.40	.000
IQ	**0.08**	0.04	**.13**	2.08	.039
SES	–0.14	0.36	–.02	–0.39	.698
Alter	0.07	0.07	.06	1.05	.296
Geschlecht[1)]	**1.94**	0.86	**.13**	2.25	.025
Förderbedarf Mathematik[2)]	**–3.15**	0.90	**–.21**	–3.48	.001
Einschränkungen im Deutschverstehen[3)]	**–1.99**	0.89	**–.13**	–2.23	.027
Gruppe$^{\text{Begl 4)}}$	–1.97	1.06	–.12	–1.85	.065
Gruppe$^{\text{Mat 4)}}$	–0.58	1.09	–.04	–0.53	.595

Korrigiertes R^2 = .34; [1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschr. im Deutschverstehen; [4)] Referenzgruppe: Gruppe$^{\text{Kontr}}$

Somit konnten Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen, die in der Gruppe^{Begl} an der unterrichtsintegrierten Mathematikförderung teilgenommen haben, ihre Mathematikleistung im Vergleich zu Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen, die einen regulären Mathematikunterricht besucht haben, nicht bedeutsam steigern, sondern sie erreichten beim Post-Test tendenziell niedrigere Mathematikleistungen im Vergleich zur Gruppe^{Kontr}. Der standardisierte Regressionskoeffizient für die Gruppe^{Mat} ist ebenfalls negativ, jedoch unterscheidet er sich nicht signifikant vom Regressionskoeffizienten für die Gruppe^{Kontr} ($\beta = -.04$, $p = .595$). Demnach weisen Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen der Gruppe^{Mat} beim Post-Test keine statistisch bedeutsame Leistungsverbesserung auf im Vergleich zu entsprechende Schülerinnen und Schülern der Gruppe^{Kontr}, die an einem regulären Mathematikunterricht teilgenommen haben, schneiden aber auch nicht signifikant schlechter ab als diese. Im Gegensatz zu den im Kapitel 6.3 präsentierten Mehrebenenmodellen, in denen alle Schülerinnen und Schüler berücksichtigt wurden, zeigt sich bei Kindern mit unterdurchschnittlichen Mathematikleistungen kein Vorteil für die Gruppe^{Mat}. In der Gruppe^{Begl} ist zudem der Unterschied zur Gruppe^{Kontr} tendenziell signifikant, wenn nur die rechenschwachen Kinder betrachtet werden.

Beim Follow-up 1 ist der standardisierte Regressionskoeffizient für die Gruppe^{Begl} zwar auch negativ, jedoch unterscheidet sich dieser Regressionskoeffizient nicht signifikant von demjenigen für die Gruppe^{Kontr} ($\beta = -.08$, $p = .250$; vgl. Tabelle 50). Für die Gruppe^{Mat} hingegen ist der standardisierte Regressionskoeffizient beim Follow-up 1 positiv, allerdings liegt auch hier kein signifikanter Unterschied im Vergleich zur Gruppe^{Kontr} vor ($\beta = .07$, $p = .283$). Beim Follow-up 1 muss für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen für beide Interventionsgruppen somit festgestellt werden, dass keine statistisch signifikante Leistungssteigerung im Vergleich zu entsprechenden Schülerinnen und Schülern aus der Kontrollgruppe festzustellen war. Für die Gruppe^{Mat} stimmen die Ergebnisse für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen somit nicht mit den Ergebnissen aus dem Kapitel 6.3 überein, in dem alle Schülerinnen und Schüler unabhängig von ihrem Leistungsniveau untersucht wurden und in dem bessere Leistungen für die Gruppe^{Mat} im Vergleich zur Gruppe^{Kontr} bestätigt werden konnten.

Tabelle 50: Multiple Regressionsanalyse für die abhängige Variable Mathematikleistung beim Follow-up 1 für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen beim Pre-Test

Variable	***B***	***SE (B)***	**β**	***t***	***p***
(Konstante)	10.07	8.53		1.18	.239
Vorwissen Mathematik	**0.86**	0.12	**.45**	7.31	.000
IQ	0.04	0.04	.06	0.99	.322
SES	–0.43	0.35	–.07	–1.21	.230
Alter	–0.03	0.07	–.02	–0.42	.674
Geschlecht[1)]	**2.35**	0.84	**.16**	2.78	.006
Förderbedarf Mathematik[2)]	**–3.17**	0.89	**–.21**	–3.58	.000
Einschränkungen im Deutsch-verstehen[3)]	**–2.29**	0.88	**–.15**	–2.61	.010
Gruppe^{Begl} [4)]	–1.19	1.03	–.08	–1.15	.250
Gruppe^{Mat} [4)]	1.15	1.07	.07	1.08	.283

Korrigiertes R^2 = .36; [1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschr. im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}

Beim Follow-up 2 zeigt sich ein ähnliches Bild wie beim Follow-up 1 (vgl. Tabelle 51). In beiden Interventionsgruppen konnten Lernende mit unterdurchschnittlichen Mathematikleistungen durch den Besuch der unterrichtsintegrierten Mathematikförderung ihre Mathematikleistungen im Vergleich zu Lernenden mit unterdurchschnittlichen Mathematikleistungen, die einen regulären Mathematikunterricht besucht haben, nicht statistisch bedeutsam steigern. Für die Gruppe^{Begl} resultiert ein negativer Regressionskoeffizient (β = –.02, p = .812), für die Gruppe^{Mat} ein positiver Regressionskoeffizient (β = .10, p = .138). Auch beim Follow-up 2 zeigen sich für die Gruppe^{Mat} abweichende Ergebnisse im Vergleich zu den Mehrebenenanalysen aus Kapitel 6.3.

Tabelle 51: Multiple Regressionsanalyse für die abhängige Variable Mathematikleistung beim Follow-up 2 für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen beim Pre-Test

Variable	***B***	***SE (B)***	**β**	***t***	***p***
(Konstante)	8.81	12.09		0.73	.467
Vorwissen Mathematik	**1.29**	0.16	**.49**	7.89	.000
IQ	0.09	0.05	.11	1.71	.089
SES	–0.75	0.48	–.09	–1.55	.123
Alter	0.05	0.10	.03	0.49	.622
Geschlecht[1)]	1.91	1.17	.09	1.64	.102
Förderbedarf Mathematik[2)]	**–4.21**	1.24	**–.20**	–3.41	.001
Einschränkungen im Deutsch-verstehen[3)]	**–3.42**	1.21	**–.16**	–2.82	.005
Gruppe^{Begl} [4)]	–0.34	1.44	–.02	–0.24	.812
Gruppe^{Mat} [4)]	2.22	1.49	.10	1.49	.138

Korrigiertes R^2 = .41; [1)] Referenzgruppe: Mädchen; [2)] Referenzgruppe: kein Förderbedarf in Mathematik; [3)] Referenzgruppe: keine Einschr. im Deutschverstehen; [4)] Referenzgruppe: Gruppe^{Kontr}

Werden nur Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen untersucht, können die Hypothesen der Fragestellungen zwei und drei, die in der vorliegenden Untersuchung für das gesamte Leistungsspektrum formuliert wurden, somit beide nicht bestätigt werden. Demnach haben weder die Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen der Gruppe^{Mat} noch der Gruppe^{Begl} ihre Mathematikleistung durch die Teilnahme an der unterrichtsintegrierten Mathematikförderung im Vergleich zu entsprechenden Schülerinnen und Schülern, die einen regulären Mathematikunterricht besucht haben, statistisch bedeutsam steigern können. Mit Blick auf die Regressionskoeffizienten der eingefügten Prädiktoren zeigen sich einige Abweichungen zu den weiter oben präsentierten Mehrebenenmodellen, die an dieser Stelle aber nicht ausgeführt werden, da in den Regressionsanalysen für die schwachen Rechnerinnen und Rechner die hierarchische Datenstruktur nicht berücksichtigt werden konnte.

7. Diskussion

In der vorliegenden Arbeit wurde der Frage nach der Wirksamkeit einer verstehensorientierten Mathematikförderung, die unterrichtsintegriert in der dritten Klasse der Primarschule umgesetzt wurde, nachgegangen und es wurden Prädiktoren auf Individual- und Klassenebene nach deren Einfluss auf die mathematische Leistung analysiert. Zunächst werden im Folgenden die drei Fragestellungen der vorliegenden Untersuchung nochmals genannt. Anschliessend werden zuerst die Ergebnisse zur zweiten und zur dritten Fragestellung bezüglich der Wirksamkeit der unterrichtsintegrierten Mathematikförderung zusammenfassend dargestellt und mit Blick auf den Forschungsstand diskutiert. Erst danach werden die zentralen Ergebnisse zur ersten Fragestellung nochmals aufgezeigt und für die einzelnen Prädiktoren auf Individual- und Klassenebene theoretisch eingebettet und diskutiert. Die Arbeit schliesst mit Hinweisen zu den Grenzen der vorliegenden Untersuchungen sowie einem Fazit, in dem Schlussfolgerungen für zukünftige Interventionsstudien gezogen werden.

Die drei Fragestellungen der vorliegenden Untersuchung lauten:

1. Fragestellung:
Welche erhobenen Merkmale auf Individual- und auf Klassenebene haben einen Einfluss auf die Mathematikleistung?

2. Fragestellung:
Kann durch eine von der Lehrperson durchgeführte, unterrichtsintegrierte Förderung im Fach Mathematik, die auf zentrale arithmetische Inhalte fokussiert und Aspekte effektiven Mathematikunterrichts berücksichtigt, im Vergleich zu einem regulären Mathematikunterricht eine Verbesserung der Mathematikleistungen erreicht werden?

3. Fragestellung:
Hat die Form, mit der die Lehrpersonen angeleitet werden (schriftliche Materialien und Anleitungen vs. intensivere Begleitung durch Fortbildungstreffen) einen Einfluss auf die Leistungsfortschritte der Schülerinnen und Schüler?

7.1 Wirksamkeit der unterrichtsintegrierten Förderung im Fach Mathematik

Bezüglich der Frage, ob Schülerinnen und Schüler der dritten Klasse der Primarschule von einer unterrichtsintegrierten Mathematikförderung profitieren können, die auf zentrale Inhalte der Arithmetik fokussiert und Aspekte effektiven Mathematikunterrichts berücksichtigt, ergibt sich in der vorliegenden Untersuchung ein Bild, das Fragen aufwirft.

7.1.1 Wirksamkeit in der Gruppe^{Mat}

Für die zweite Fragestellung kann die Hypothese, wonach Schülerinnen und Schüler im Vergleich zu einem regulären Mathematikunterricht von einer unterrichtsintegrierten Mathematikförderung profitieren, wie sie in der vorliegenden Untersuchung umgesetzt wurde, für die Gruppe^{Mat} für alle Messzeitpunkte sowohl in der Stichprobe^{alle} als auch in der Stichprobe^{IQstreng} bestätigt werden (vgl. Kap. 6.3). Demnach weisen die Schülerinnen und Schüler der Gruppe^{Mat} im Vergleich zur Gruppe^{Kontr} beim Post-Test eine Mathematikleistung auf, die im Durchschnitt um 1.51 Punkte (Stichprobe^{alle}) bzw. um 1.50 Punkte (Stichprobe^{IQstreng}) höher liegt, bei einem Test, in dem insgesamt 41 Punkte erzielt werden konnten. Beim Follow-up 1 beträgt dieser Unterschied 1.71 (Stichprobe^{alle}) bzw. 1.54 Punkte (Stichprobe^{IQstreng}) bei gleichem Punktemaximum. Beim Follow-up 2 ergibt sich ein Unterschied von 1.92 Punkten (Stichprobe^{alle}) bzw. von 1.88 Punkten (Stichprobe^{IQstreng}), bei einem Test mit 60 Punkten. Insbesondere beim Follow-up 1 fällt auf, dass für die Stichprobe^{alle} ein höherer Regressionskoeffizient resultiert als für die Stichprobe^{IQstreng}. Werden demnach auch ältere Schülerinnen und Schüler für die Analysen berücksichtigt, die im Durchschnitt über weniger mathematisches Vorwissen und einen niedrigeren IQ verfügen als die Schülerinnen und Schüler der Gesamtstichprobe (vgl. Kap. 5.5.1), fällt der Unterschied zwischen der Gruppe^{Mat} und der Gruppe^{Kontr} höher aus. Dies widerspiegelt sich auch in den Effektstärken.

Die Effektstärke für die Gruppe^{Mat} beträgt in der Stichprobe^{alle} beim Post-Test 0.29, beim Follow-up 1 0.40 und beim Follow-up 2 0.32 sowie für die Stichprobe^{IQstreng} beim Post-Test 0.29, beim Follow-up 1 0.33 und beim Follow-up 2 0.30. Gemäss Cohen (1988) sind diese Effektstärken alle als klein einzuschätzen (mittlere Effektstärken werden ab 0.50 angenommen). Allerdings muss bei Interventionsstudien im natürlichen Kontext der Schule angenommen werden, dass viele verschiedene Faktoren einen Einfluss auf die Ergebnisse haben, sodass die Erklärungskraft eines einzelnen Prädiktors per se begrenzt ist (Fuchs et al., 2012; Trautwein et al., 2009; vgl. Kap. 5.6.2). Es ist daher wichtig, die gefundenen Effektstärken denjenigen aus anderen vergleichbaren Studien gegenüber zu stellen. Dabei fällt die grosse Spannbreite der gefundenen Effektstärken auf (vgl. Kap. 3.1). So berichten Woodward und Brown (2006) von Effektstärken von 1.23 (Rechenfertigkeiten) bzw. 1.95 (konzeptuelles Verständnis; vgl. Kap. 3.1.4). In dieser Studie haben jedoch nur zwei Lehrpersonen die mathematische Förderung in enger Begleitung durch die Forschungsgruppe umgesetzt. Im Gegensatz zur vorliegenden Untersuchung muss daher wenig Variabilität bei der Umsetzung der Förderung angenommen werden. Zusätzlich ist zu bedenken, dass bei Woodward und Brown (2006) zwar ebenfalls eine Förderung im Rahmen eines regulären Klassenunterrichts untersucht wurde, der jedoch nur von lernschwachen Schülerinnen und Schülern besucht wurde. Die Lehrpersonen mussten in dieser Interventionsstudie somit ihr unterrichtliches Handeln vermutlich nicht an ein so breites Leistungsspektrum, wie dies in einer Regelklasse üblicherweise der Fall ist, anpassen, sondern konnten sich stärker auf die Bedürfnisse von schwachen Rechnerinnen und Rechner konzentrieren. Sowohl die geringe Variabilität bei der Umsetzung als auch die Zusammensetzung der Stichprobe könnten somit dazu beigetragen haben, dass in der genannten Studie trotz stark natürlichem Kontext sehr hohe Effektstärken im Vergleich zur vorliegenden Untersuchung für die mathematische Förderung gefunden wurden.

Im Vergleich dazu bewegen sich die von Moser Opitz et al. (2017) gefundenen Effektstärken in einer ähnlicheren Grössenordnung wie in der hier vorliegenden Arbeit. So konnten für das teilweise klassenintegrierte Fördersetting Effektstärken von 0.32 (Post-Test) bzw. 0.48 (Follow-up) für den Gesamttest belegt werden (vgl. Kap. 3.1.4). Auffallend ist dabei, dass sich auch in dieser Studie für die Follow-up-Messungen höhere Werte ergaben als für den Post-Test, was darauf hinweist, dass sich bei einer verstehensorientierten Mathematikförderung Fördereffekte insbesondere auch mittel- und längerfristig positiv auswirken (vgl. hierzu auch Sinner, 2011; Kap. 3.1.2). In Studien, in denen mathematische Vorläuferfertigkeiten im Sinne des Entwicklungsmodells von Krajewski (2003, 2005, 2007) in Kleingruppen trainiert wurden, zeigten sich bezüglich Leistungen in standardisierten Mathematiktests, die Rechenaufgaben beinhaltet haben, Effektstärken von 0.37 bis 0.77, wobei diese als Transfereffekt interpretiert wurden (Ennemoser & Krajewski, 2007; Ennemoser et al., 2015; Sinner, 2011; vgl. Kap. 3.1.2). In den Studien von Bryant et al. (2008) und Wißmann et al. (2013) wurden hingegen keine Effektstärken dokumentiert, so dass kein Vergleich möglich ist (vgl. Kap. 3.1.3). Zu vorherigen Ausführungen muss jedoch angemerkt werden, dass in allen genannten Untersuchungen jeweils nur Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen an der Förderung teilgenommen haben, was einen Einfluss auf die Höhe der gefundenen Effektstärken haben könnte. Trotz Empfehlungen von Cohen (1988) können aufgrund der oben genannten Ergebnisse aus vergleichbaren Interventionsstudien und der Tatsache, dass die hier untersuchte unterrichtsintegrierte Mathematikförderung in einem sehr realen bzw. natürlichen schulischen Kontext unter Berücksichtigung eines breiten Leistungsspektrums stattfand, die in dieser Arbeit gefundenen Effekte als bedeutsam eingeschätzt werden (vgl. Fuchs et al., 2012; Trautwein et al., 2009). Dies wird insbesondere durch die von Moser Opitz et al. (2017) dargelegten Effektstärken unterstützt, die sich ebenfalls in einem stark natürlichen Kontext gezeigt haben.

Für die Gruppe^{Mat} kann somit die Hypothese der zweiten Fragestellung bestätigt werden. Demnach kann auf Basis der Ergebnisse für die Gruppe^{Mat} davon ausgegangen werden, dass eine unterrichtsintegrierte Mathematikförderung wirksam ist, in der

- auf zentrale mathematische Inhalte fokussiert wird, die schwachen Rechnerinnen und Rechnern Mühe bereiten (vgl. Kap. 2.4, 3.3.1 und 4.2.1),
- geeignete Veranschaulichungen und Arbeitsmittel verwendet werden (vgl. Kap. 3.3.1 und 4.2.2),
- produktive Übungsformate zum Einsatz kommen (vgl. Kap. 3.3.1 und 4.2.3),
- gemeinsame Lernsituationen geschaffen werden (vgl. Kap. 3.2.2 und 4.2.4),
- eine adaptive Lernbegleitung umgesetzt wird (vgl. Kap. 3.2.1 und 4.2.5).

Trotz dieser Ergebnisse für die Gruppe^{Mat}, die im Einklang mit Forschungsergebnissen zu mathematischen Förderansätzen stehen (vgl. Kap. 3.1), müssen für die vorliegende Untersuchung einige Einschränkungen beachtet werden, die im Folgenden diskutiert werden.

7.1.2 Wirksamkeit in der Gruppe^{Begl}

In der Gruppe^{Begl} wurde die gleiche unterrichtsintegrierte Mathematikförderung umgesetzt wie in der Gruppe^{Mat}. Die Lehrpersonen der Gruppe^{Begl} haben jedoch eine intensivere Begleitung in Form von zusätzlichen Professionalisierungsmassnahmen erhalten (vgl. Kap. 5.3.2). So wurde den Lehrpersonen in drei Fortbildungstreffen u.a. vertieftes Wissen zu Rechenschwäche vermittelt und Reflexionsgelegenheiten basierend auf Unterrichtsvideos und den eigenen Erfahrungen bei der Umsetzung der Förderung ermöglicht (vgl. Kap. 3.4). Die Ergebnisse der vorliegenden Arbeit zeigen jedoch, dass die Schülerinnen und Schüler der Gruppe^{Begl} bei keinem Messzeitpunkt bessere Mathematikleistungen erzielt haben als die Gruppe^{Kontr}. Zudem schnitt die Gruppe^{Begl} beim Post-Test und beim Follow-up 2 in beiden Untersuchungsstichproben signifikant schlechter ab als die Gruppe^{Mat} (vgl. Kap. 6.3). Demnach muss die Hypothese der zweiten Fragestellung für die Gruppe^{Begl} abgelehnt werden. Trotz entsprechender Ergebnisse für die Gruppe^{Mat} kann daher nicht grundsätzlich angenommen werden, dass eine unterrichtsintegrierte Mathematikförderung, die auf zentrale Inhalte fokussiert und Aspekte effektiven Mathematikunterrichts berücksichtigt, wirksam ist. Zudem muss auch die Hypothese für die dritte Fragestellung, wonach eine intensivere Begleitung der Lehrpersonen durch Fortbildungstreffen zu einem zusätzlichen Lernzuwachs bei den Schülerinnen und Schülern der Gruppe^{Begl} im Vergleich zur Gruppe^{Mat} führen sollte, abgelehnt werden.

Dass die dritte Fragestellung abgelehnt werden muss, ist erklärungsbedürftig. So zeigen Forschungsergebnisse, dass Professionalisierungsmassnahmen, wie sie im Rahmen der intensiveren Begleitung der Lehrpersonen in Gruppe^{Begl} umgesetzt wurden (vgl. Kap. 5.3.2), sowohl die professionellen Kompetenzen und das professionelle Wissen der Lehrpersonen als auch die Leistungen der Schülerinnen und Schüler steigern können (Carpenter et al., 1989; Kunter et al., 2011; Lipowsky, 2011; Staub, 2014). Allerdings handelt es sich in der vorliegenden Untersuchung um eine Professionalisierungsmassnahme von kurzer Dauer (dreimal 2.5 bis 3 Stunden). Die Effektivität sehr kurzer Professionalisierungsmassnahmen wird jedoch eher kritisch gesehen. Dies bestätigen Ergebnisse einer Metaanalyse aus den USA, wonach keine signifikanten Effekte für Professionalisierungsmassnahmen gefunden wurden, die 14 oder weniger Stunden umfassen (Yoon, Duncan, Lee, Scarloss & Shapley, 2007). Die Ergebnisse der vorliegenden Untersuchung weisen somit darauf hin, dass der gewählte zeitliche Umfang für die Professionalisierungsmassnahmen wohl nicht ausreichend war. Aufgrund der schon hohen Belastung der Lehrpersonen im Rahmen der Untersuchung war es jedoch nicht möglich, eine zeitlich aufwendigere Begleitung durchzuführen. Für zukünftige Studien wäre es deshalb interessant, eine intensivere Begleitung mit umfangreicheren Professionalisierungsmassnahmen bei der Implementierung von Fördermassnahmen zu untersuchen. Allerdings muss hierzu beachtet werden, dass dies die Bereitschaft von Lehrpersonen zur Teilnahme an Studien negativ beeinflussen kann. Im Folgenden werden weitere Gründe für die unterschiedlichen Ergebnisse in den beiden Interventionsgruppen diskutiert.

7.1.3 Mögliche Gründe für die unterschiedlichen Ergebnisse in den Interventionsgruppen

Stichprobenspezifische Merkmale als Gründe für die unterschiedlichen Ergebnisse

Wie die Mehrebenenmodelle im Kapitel 6.4 zeigen, gibt es zwei stichprobengruppenbezogene Faktoren, die zu den unterschiedlichen Ergebnissen bezüglich der Wirksamkeit der Förderung in den beiden Interventionsgruppen beigetragen haben könnten. Gemäss Mehrebenenmodell 13 weisen in der Gruppe^{Begl} und in der Gruppe^{Kontr} ältere Schülerinnen und Schüler schlechtere Mathematikleistungen auf als jüngere Schülerinnen und Schüler (vgl. Kap. 6.4.1.). In der Gruppe^{Mat}, in der die Förderung wirksam war, ist dies nicht der Fall. Die Schülerinnen und Schüler in der Gruppe^{Mat} sind zudem im Vergleich zur Gruppe^{Kontr} signifikant bzw. im Vergleich zur Gruppe^{Begl} tendenziell älter (vgl. Kap. 5.5.1). Auf Basis der Altersstruktur der Gesamtstichprobe kann angenommen werden, dass es sich bei einem grossen Teil der älteren Schülerinnen und Schüler um Kinder handelt, die eine Klasse repetiert oder aufgrund einer späteren Einschulung länger den Kindergarten besucht haben bzw. während zweier Jahre in einer Einführungsklasse beschult wurden und somit im Verlaufe ihrer bisherigen Schulzeit mit grosser Wahrscheinlichkeit schulische Schwierigkeiten gezeigt haben. Eine mögliche Hypothese wäre daher, dass diese Schülerinnen und Schüler durch den längeren Schulbesuch sowie die Gründe, die hierzu geführt haben, spezifische Voraussetzungen aufweisen, die dazu beigetragen haben, dass sie stärker auf die Förderung angesprochen haben. Da die Gruppe^{Mat} den höchsten Altersdurchschnitt der drei Stichprobengruppen aufweist, wäre dieser Effekt daher in dieser Gruppe am stärksten ausgeprägt. Eine mögliche Erklärung für die abweichenden Ergebnisse bezüglich der Wirksamkeit der Förderung in den beiden Interventionsgruppen würde daher lauten, dass bei einem hohen Anteil älterer Schülerinnen und Schüler die hier untersuchte unterrichtsintegrierte Mathematikförderung aufgrund der spezifischen Voraussetzungen dieser Personengruppe wirksamer ist, was zu den positiven Ergebnissen für die Gruppe^{Mat} beigetragen hat.

Im Mehrebenenmodell 14 hat sich des Weiteren gezeigt, dass sich in der Gruppe^{Begl} ein hoher Anteil an Schülerinnen und Schülern mit Einschränkungen im Deutschverstehen in den Klassen signifikant negativ auf die individuelle Mathematikleistung ausgewirkt hat, nicht jedoch in der Gruppe^{Mat} und der Gruppe^{Kontr} (vgl. Kap. 6.4.2). Hierzu muss berücksichtigt werden, dass die Gruppe^{Begl}, in der die unterrichtsintegrierte Förderung nicht wirksam war, einen tendenziell höheren Anteil an Schülerinnen und Schülern mit Einschränkungen im Deutschverstehen aufweist als die anderen beiden Stichprobengruppen (vgl. Kap. 5.5.1). Eine mögliche Hypothese lautet daher, dass bei einem sehr hohen Anteil an Schülerinnen und Schülern mit Einschränkungen im Deutschverstehen in den Klassen zusätzliche bzw. andere Fördermassnahmen notwendig sind, als in der hier untersuchten unterrichtsintegrierten Förderung umgesetzt wurden. Gemäss aktuellem Forschungsstand (vgl. Kap. 2.3.3) wird angenommen, dass sprachliche Schwierigkeiten insbesondere auch den Erwerb von konzeptuellem Wissen und prozeduralen Kompetenzen erschweren (Paetsch et al., 2015; Prediger et al., 2015; Vukovic & Lesaux, 2013). Angesichts der starken Gewichtung von Versprachlichungen eigener Rechen- und Denkwege sowie von gemeinsamen Reflexionsrunden im Klassenplenum im Rahmen der vorliegenden unterrichtsintegrierten Mathematikförderung, scheint es plausibel, dass es zu entsprechenden Zusammenhängen gekommen ist. Für zukünftige Untersuchungen wäre es daher wichtig, das komplexe Bedingungsgefüge zwischen verschiedenen sprachlichen Aspekten (z.B. mathematisch-fachsprachliche Kenntnisse, allgemeines Sprachverständnis etc.; vgl. Bochnik & Ufer, 2016) und den mathematischen Leistungen unter Berücksichtigung weiterer individueller Faktoren (z.B. SES und IQ) auch im Rahmen von Interventionsstudien zur mathematischen Förderung detaillierter zu untersuchen. Zu beiden oben genannten möglichen Gründen für die unterschiedlichen Ergebnisse in den beiden Interventionsgruppen muss jedoch angemerkt werden, dass sich die Effekte für das Alter und den Anteil Kinder mit Einschränkungen im Deutschverstehen jeweils nur bei einem einzelnen Messzeitpunkt gezeigt haben. Daher können diese nicht grundsätzlich erklären, weshalb die Förderung in der Gruppe$^{\text{Begl}}$ nicht effektiv war. Es kann jedoch als plausibel angenommen werden, dass diese Faktoren zu den unterschiedlichen Ergebnissen beigetragen haben.

Qualitative Unterschiede bei der Umsetzung als Grund für die unterschiedlichen Ergebnisse
Eine weitere Möglichkeit ist, dass die Qualität, mit der die unterrichtsintegrierte Mathematikförderung von den jeweiligen Lehrpersonen umgesetzt wurde, Ursache für die gefundenen Gruppenunterschiede bezüglich Wirksamkeit ist. Zwar wurde in der vorliegenden Untersuchung durch detaillierte Anleitungen vorgegeben, wie die Förderung von den Lehrpersonen im Unterricht umgesetzt werden soll (vgl. Kap. 5.3.1). Allerdings können qualitativ hochstehende Lehrprozesse dadurch nicht garantiert werden, da die Qualität der Lehrprozesse von weiteren Lehrpersonenmerkmalen beeinflusst wird (z.B. professionelle Kompetenzen, Überzeugungen, Lehrerfahrungen; Fend, 1981; Helmke & Weinert, 1997; Seidel, 2014; vgl. Kap. 4.2.6). Interessante Hinweise hierzu bietet eine Dissertation, die im gleichen Projekt wie die vorliegende Arbeit entstanden ist (Pfister, 2016). Mithilfe von Videoanalysen und einem hochinferenten Rating konnte dort aufgezeigt werden, dass die Adaptivität der Lernbegleitung zwischen den einzelnen Lehrpersonen der Interventionsgruppen teilweise sehr unterschiedlich ausfiel (Pfister, 2016).[29] Insbesondere im Bereich der Diskursanregung ergaben die Analysen, dass die Hälfte der untersuchten Lehrpersonen über eine gering bis eher gering ausgeprägte Adaptivität der diesbezüglichen Lernbegleitung gezeigt hat (Pfister, 2016). Im vorliegenden Konzept einer unterrichtsintegrierten Förderung im Mathematikunterricht kommt jedoch gerade der Diskursanregung ein wichtiger Stellenwert zu, da die sprachliche Begleitung von Lernprozessen sowie die durch die Lehrperson initiierte gemeinsame Reflexion im Klassenverband zentrale Bestandteile der Förderung waren. Dass eine inhaltlich fast identische Förderung bei unterschiedlichen Umsetzungsformen nicht zu vergleichbaren Ergebnissen bezüglich der Effektivität der Förderung führen muss, hat sich zudem auch in einer anderen Untersuchung gezeigt (Freesemann, 2014; Moser Opitz et al., 2017; vgl. Kap. 3.1.4).

Mit Blick auf die Einordnung der unterrichtsintegrierten Mathematikförderung im Angebots-Nutzungs-Modell (Fend, 1981; Helmke & Weinert, 1997; Seidel, 2014), kann daher vermutet werden, dass durch die schriftlichen Anleitungen und Materialien, die den Lehrpersonen für die hier untersuchte unterrichtsintegrierte Mathematikförderung zur Verfügung gestellt wurden, die

[29] Für die Kontrollklassen liegen keine entsprechenden Daten vor (vgl. Kap. 5.2).

Kompetenzen der Lehrpersonen und die Lehrprozesse im Unterricht zwar teilweise gesteigert werden konnten, was sich positiv auf die Lernaktivitäten und im Falle der Gruppe^{Mat} auch positiv auf die Leistungen der Schülerinnen und Schüler in Mathematik ausgewirkt hat (vgl. weisse Pfeile in Abbildung 7 im Kap. 4.2.6). Allerdings muss auf Basis der vorliegenden Ergebnisse angenommen werden, dass insbesondere im Bereich der Lehrprozesse die schriftlichen Anleitungen und Materialien nicht in allen Klassen und nicht bei allen Lehrpersonen dazu geführt haben, dass die Lehrprozesse im Unterricht qualitativ so stark gesteigert werden konnten, dass die Lernaktivitäten der Schülerinnen und Schüler sich ebenfalls qualitativ verbessert haben, und sich dies positiv auf die Leistungen in Mathematik ausgewirkt hat (vgl. Brühwiler & Blatchford, 2011). Unter zusätzlicher Berücksichtigung der weiter oben diskutierten stichprobenspezifischen Merkmale, bei denen angenommen werden kann, dass sie einen Einfluss auf die unterschiedliche Wirksamkeit der unterrichtsintegrierten Förderung gehabt haben, kann daher vermutet werden, dass qualitative Unterschiede bei der Umsetzung zwischen den Lehrpersonen ebenfalls zu den diskrepanten Ergebnissen in den beiden Interventionsgruppen beigetragen haben.

Hinsichtlich der Umsetzungsqualität sind auch Rekontextualisierungsmechanismen im Sinne von Fend (2008) als mögliche Erklärung für abweichende Lehr- und Lernprozesse in den verschiedenen Klassen in Betracht zu ziehen. Demnach haben verschiedene Faktoren auf der Mikro-, Meso- und Makroebene einen Einfluss auf die Umsetzung von Vorgaben, die von aussen an die Schule herangetragen werden (Fend, 2008). Bei der Implementierung von Förderprogrammen muss daher damit gerechnet werden, dass Lehrpersonen bei der Umsetzung im eigenen Unterricht klassenbezogene und schulhausspezifische Konstellationen sowie persönliche Voraussetzungen und Erwartungshaltungen berücksichtigen und die Förderung entsprechend adaptieren. Im Kontext von Interventionsstudien bedeutet dies, dass von „[…] empirischen Variationen des faktischen operativen Handelns“ ausgegangen werden muss (Fend, 2008, S. 27). Mit Blick auf die erwartungswidrigen Ergebnisse der Wirksamkeit der Förderung in der Gruppe^{Begl} muss in diesem Zusammenhang auch bedacht werden, dass der Besuch der Fortbildungstreffen bei den Lehrpersonen dieser Stichprobengruppe zu einer weniger autonomen Rekontextualisierung der schriftlichen Materialien und Unterlagen geführt haben könnte, so dass eine weniger stark ausgeprägte Adaption der Förderung an lehrpersonenbezogene und klassenspezifische Faktoren erfolgt ist. Dies könnte sich wiederum negativ auf die mathematische Leistungsentwicklung der Lernenden in der Gruppe^{Begl} ausgewirkt haben. Für zukünftige Untersuchungen wäre es aus den genannten Gründen wichtig, die allgemeine Umsetzungsqualität sowie die Qualität der jeweiligen Lehrprozesse in allen Klassen und über einen möglichst grossen Zeitraum systematisch zu erfassen, um so deren Auswirkungen auf die mathematischen Leistungen der Schülerinnen und Schüler und die Wirksamkeit der Förderung analysieren zu können.

7.1.4 Ergebnisse der Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen

Da die Wirksamkeit der unterrichtsintegrierten Mathematikförderung sich in der vorliegenden Untersuchung nur für eine Interventionsgruppe gezeigt hat, wurden zusätzlich zu den Mehrebenenanalysen, in denen alle Schülerinnen und Schüler einer Klasse unabhängig ihres Leistungsniveaus untersucht wurden, auch Analysen durchgeführt, in denen nur Schülerinnen und Schüler des untersten Leistungsquartils berücksichtigt wurden. Für diese Schülerinnen und Schüler zeigt sich für alle drei Messzeitpunkte weder in Gruppe^{Mat} noch der Gruppe^{Begl} ein signifikanter Vorteil für den Besuch der unterrichtsintegrierten Mathematikförderung im Vergleich zu einem regulären Mathematikunterricht. Werden die Fragestellungen zwei und drei somit auf die Analyse der Mathematikleistungen von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen im Pre-Test reduziert, kann weder die zweite noch die dritte Hypothese der vorliegenden Untersuchung bestätigt werden. Angesichts verschiedener effektiver Interventionsstudien zur Förderung bei Rechenschwäche (z.B. Bryant et al., 2008; Moser Opitz et al., 2017; Wißmann et al., 2013; Woodward & Brown, 2006) entsprechen die nicht erfolgreichen Ergebnisse für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen im Rahmen der hier untersuchten Mathematikförderung somit nicht den Erwartungen. In den genannten Interventionsstudien fand ebenfalls eine verstehensorientierte Förderung statt, in der auf zentrale Inhalte der Mathematik fokussiert, produktive Übungsformate eingesetzt und geeignete Veranschaulichungen und Arbeitsmittel verwendet wurden (vgl. Kap. 3.1.3 und 3.1.4). Es stellt sich daher die Frage, wieso diese ansonsten als wirksam erwiesenen Förderaspekte in der vorliegenden Untersuchung nicht den erwarteten Effekt bei Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen bewirkt haben.

Hierzu muss beachtet werden, dass in den zitierten Interventionsstudien die Förderung in Unterrichtssettings umgesetzt wurde, die von demjenigen, das in der hier untersuchten Studie zum Einsatz kam, abgewichen sind. So wurde die Förderung in den meisten Untersuchungen in Kleingruppen mit rechenschwachen Schülerinnen und Schülern und nicht unterrichtsintegriert durchgeführt (Bryant et al., 2008; Bryant et al., 2011; Ennemoser & Krajewski, 2007; Ennemoser et al., 2015; Sinner, 2011; Wißmann et al., 2013). Bei einer mathematischen Förderung in Kleingruppen kann im Vergleich zu einer Förderung im Unterricht sowohl hinsichtlich der inhaltlichen Gestaltung als auch hinsichtlich der Lernbegleitung von abweichenden Anforderungen ausgegangen werden. So muss sowohl bei der Planung als auch bei der Umsetzung der Förderung in Kleingruppen ein weniger breites mathematisches Leistungsspektrum berücksichtigt werden. Es kann daher angenommen werden, dass dies in den genannten Studien dazu beigetragen hat, dass im Rahmen der jeweiligen mathematischen Förderung besser auf die individuellen Bedürfnisse der Lernenden eingegangen und somit eine optimalere Passung zwischen unterrichtlichem Angebot und individuellen Lernvoraussetzungen geschaffen werden konnte. Dies könnte sich in den referierten Studien positiv auf die Leistungsentwicklung der Schülerinnen und Schüler mit schwachen Mathematikleistungen ausgewirkt haben (vgl. Parsons et al., 2018).

Des Weiteren könnte die Auswahl der Probandinnen und Probanden mit Blick auf die abweichenden Ergebnisse ebenfalls eine Rolle spielen. So fand bei Woodward und Brown (2006) eine inhaltlich mit der eigenen Studie vergleichbare Förderung im Rahmen von regulären Mathematiklektionen statt, allerdings handelte es sich in dieser Studie bei allen Schülerinnen und Schülern um Risikokinder hinsichtlich Rechenschwäche. In der Studie von Freesemann (2014) und Moser Opitz et al. (2017) wurde eine zumindest teilweise klassenintegrierte Förderung umgesetzt, allerdings haben auch dort nur rechenschwache Schülerinnen und Schüler an der Förderung teilgenommen. Zudem wurde zusätzlich zum Klassensetting (in dem die Schülerinnen und Schüler während zweier regulärer Mathematiklektionen individuell an den Inhalten der Förderung arbeiten konnten) auch während einer Lektion pro Woche eine individuelle Begleitung bei der Aufgabenauswahl angeboten. In der hier untersuchten unterrichtsintegrierten Mathematikförderung mussten die Lehrpersonen hingegen während der gesamten Interventionszeit sowohl leistungsschwache als auch leistungsstarke Schülerinnen und Schüler im Rahmen der regulären Mathematiklektionen gemeinsam fördern. Es stand im Vergleich zu anderen Studien somit weniger Zeit für die individuelle Begleitung der Lernenden mit unterdurchschnittlichen Mathematikleistungen zur Verfügung.

Aufgrund der genannten Überlegungen stellt sich daher die Frage, ob unter Berücksichtigung von Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen

a) es in einem unterrichtsintegrierten Fördersetting noch stärker Individualisierungs- und Differenzierungsmassnahmen braucht, allenfalls auch in Form von punktueller Einzel- oder Kleingruppenförderung,

b) die Lernbegleitung durch die Lehrpersonen adaptiver und individueller gestaltet werden müsste, so dass eine stärkere Passung zwischen Fördermassnahmen und individuellen Lernvoraussetzungen erzielt wird,

c) oder ob ein unterrichtsintegriertes Fördersetting für Lernende mit unterdurchschnittlichen Mathematikleistungen an sich nicht ausreichend ist.

Auf diese Fragen kann mit Hilfe der zur Verfügung stehenden Daten aus der vorliegenden Untersuchung keine abschliessende Antwort gegeben werden. Dennoch sollen diese Punkte mit Blick auf die in der vorliegenden Arbeit getroffenen didaktischen und fachdidaktischen Massnahmen kurz diskutiert werden.

Angesichts der Bestrebung in der vorliegenden Untersuchung eine unterrichtsintegrierte Mathematikförderung umzusetzen, wurden auf Basis verschiedener Argumente (Feuser, 2013; Freudenthal, 1974; Garrote et al., 2017; Stöckli et al., 2014; Wiener & Tardif, 2004; vgl. Kap. 3.2.2) bewusst gemeinsame Lernsituationen geschaffen und Unterrichtsphasen eingebaut, in denen alle Schülerinnen und Schüler zusammen im Klassenplenum mathematische Konzepte und Prozeduren erarbeitet (Einführungsphasen) und im fachlichen Austausch miteinander ihre Erkenntnisse diskutiert und reflektiert haben (Reflexionsphase). Ergänzt wurden diese gemeinsamen Phasen durch individuelle Arbeitsphasen, in denen die Schülerinnen und Schüler auf unterschiedlichen Lernniveaus das Gelernte vertiefen und individuelle Lücken im Lernstoff schliessen konnten (vgl. Kap. 5.3.1). Für die individuellen Arbeitsphasen wurden dafür Individualisierungs- und Differenzierungsmassnahmen gemäss aktuellem Stand des Wissenschaftsdiskurses der Mathematikdidaktik eingesetzt. Dazu gehörten die Fokussierung auf zentrale Inhalte der Mathematik und deren Einführung bzw. Thematisierung auf Basis zentraler mathematikdidaktischer Erkenntnisse, die Verwendung operativ strukturierter Übungsformate unter Berücksichtigung unterschiedlicher Leistungsniveaus sowie der Einsatz von geeigneten Veranschaulichungen und Arbeitsmitteln (z.B. Freesemann, 2014; Häsel-Weide, 2016; Krauthausen & Scherer, 2007; Scherer & Moser Opitz, 2010; W. Schneider et al., 2013; Van de Walle, 2007; Wartha & Schulz, 2011; Wittmann, 1992; vgl. Kap. 3.3). Angesichts dessen, dass sich solche Massnahmen in Interventionsstudien für rechenschwache Schülerinnen und Schüler als wirksam erwiesen haben (z.B. Bryant et al., 2011; Ennemoser & Krajewski, 2007; Moser Opitz et al., 2017; Wißmann et al., 2013; Woodward & Brown, 2006), werden diese Massnahmen aufgrund der in dieser Arbeit gefundenen Ergebnisse für Lernende mit unterdurchschnittlichen Mathematikleistungen nicht grundsätzlich in Frage gestellt. Zu überprüfen wäre hingegen für zukünftige Untersuchungen mit einem unterrichtsintegrierten Setting, ob gemeinsame Lernsituationen, wie sie hier zum Einsatz kamen, zeitlich reduziert werden sollten zugunsten von stärker individualisierten Einführungs- und Reflexionssequenzen z.B. in Kleingruppen unter Einbezug einer Schulischen Heilpädagogin bzw. eines Schulischen Heilpädagogen. Ebenso müsste geprüft werden, ob Individualisierungs- und Differenzierungsmassnahmen auch in den Arbeitsphasen noch verstärkter zum Einsatz kommen müssten bzw., ob die Gestaltung der Arbeitsphasen noch stärker an den individuellen Bedürfnissen der einzelnen Schülerinnen und Schüler angepasst werden müsste. Z.B. könnte durch regelmässige Standortbestimmungen noch stärker eine Passung zwischen dem jeweiligen Unterrichtsangebot und den individuellen Lernvoraussetzungen (Vygotsky, 1978) erzielt werden. In der vorliegenden Untersuchung fand eine Standortbestimmung nur zu Beginn der drei Unterrichtseinheiten statt (vgl. Kap. 5.3.1).

In der hier untersuchten Mathematikförderung wurden des Weiteren Massnahmen für eine adaptive Lernbegleitung getroffen, die die individuelle Passung zwischen den Lernvoraussetzungen und dem Förderangebot zusätzlich unterstützen, und zudem dabei helfen sollten, die Lernprozesse zu strukturieren. Aufgrund der fehlenden Wirksamkeit der Förderung für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen muss jedoch vermutet werden, dass die Lernbegleitung insbesondere für diese Lernenden nicht ausreichend adaptiv und individuell gestaltet war. Hinweise, die diese Vermutung stützen, finden sich in den Ergebnissen der Videoanalysen von Pfister (2016), die ebenfalls Teil des Gesamtprojektes waren. Es zeigte sich, dass etwa 40% der Lehrpersonen der beiden Interventionsgruppen eine geringe bis eher geringe Ausprägung der adaptiven Lernbegleitung im Bereich der kognitiven Aktivierung und im Umgang mit Fehlern erreicht haben und für die Diskursanregung dies sogar auf etwa 50% der Lehrpersonen zutraf (Pfister, 2016). Da gerade für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen in heterogenen Lerngruppen als unabdingbar angesehen wird, dass sie eine adaptive, individuelle Lernbegleitung erhalten (vgl. Kap. 3.2.1), könnte dies zu den erwartungswidrigen Ergebnissen für diese Personengruppe beigetragen haben. Es wäre daher für zukünftige Interventionsstudien wichtig zu prüfen, mit welchen Massnahmen erreicht werden kann, dass die Lehrpersonen die Lernbegleitung adaptiver und individueller gestalten. Auch hier könnten regelmässigere und ausführlichere Standortbestimmungen dazu beitragen, dass die Lernbegleitung stärker an den Bedürfnissen der Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen ansetzt, und so eine bessere Passung zwischen individuellen Lernvoraussetzungen und Fördermassnahmen erreicht wird.

Die aufgrund der Ergebnisse der vorliegenden Arbeit für Lernende mit unterdurchschnittlichen Mathematikleistungen in Erwägung zu ziehende Vermutung, dass ein unterrichtsintegriertes

Setting für eine verstehensorientierte Mathematikförderung für diese Personengruppe in der Primarschule an sich nicht ausreichend ist, findet Unterstützung durch die Ergebnisse der Metaanalyse von Ise et al. (2012), wonach rechenschwache Schülerinnen und Schüler besonders stark von einer Einzelförderung profitieren, eine Gruppenförderung sowie eine klassenweise Förderung hingegen deutlich weniger effektiv ist (vgl. Kap. 3.1.1). Da es jedoch auch Studien gibt, in denen eine teilweise klassenintegrierte Förderung (Freesemann, 2014; Moser Opitz et al., 2017) bzw. eine klassenintegrierte Förderung im Rahmen eines Unterrichts nur für lernschwache Schülerinnen und Schüler (Woodward & Brown, 2006) erfolgreich umgesetzt werden konnte, kann ein solches Fördersetting auf Basis der in dieser Arbeit gefundenen Ergebnisse nicht grundsätzlich als unzureichend beurteilt werden. Eine abschliessende Einschätzung, ob eine unterrichtsintegrierte Umsetzung einer Mathematikförderung für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen geeignet ist oder nicht, wird daher aufgrund der in dieser Arbeit gefundenen Ergebnisse und unter Berücksichtigung erfolgreicher Interventionsstudien mit zumindest teilweisem Klassensetting als verfrüht eingeschätzt, und eine nach wie vor unklare Forschungslage angenommen. Es wird jedoch empfohlen dieser Frage in weiteren Forschungsprojekten nachzugehen. Was hinsichtlich zukünftiger Interventionsstudien dabei beachtet werden sollte, wird im Rahmen des abschliessenden Fazits noch im Detail aufgezeigt (vgl. Kap. 7.4).

7.2 Prädiktoren für die Mathematikleistung auf Individual- und Klassenebene

Im Folgenden werden nun die Ergebnisse zur ersten Fragestellung zusammenfassend präsentiert und diskutiert. Dabei geht es um die Frage nach Variablen auf Individual- und Klassenebene, für die sich im Rahmen der vorliegenden Untersuchung gezeigt hat, dass sie einen Einfluss auf die Mathematikleistung haben. Zunächst wird ein Überblick gegeben, danach werden die Ergebnisse für die einzelnen Variablen diskutiert und theoretisch eingeordnet.

Auf *Individualebene* haben sich bei allen Messzeitpunkten (Post-Test, Follow-up 1 und Follow-up 2) und in beiden Untersuchungsstichproben (Stichprobe^{alle} und Stichprobe^{IQstreng}) die folgenden Variablen als signifikante Prädiktoren für die Mathematikleistung erwiesen:

- Mathematisches Vorwissen
- IQ
- Geschlecht
- Förderbedarf in Mathematik

Es kann davon ausgegangen werden, dass diese Variablen wichtige Determinanten für die Mathematikleistung von Schülerinnen und Schülern in der dritten und der vierten Klasse der Primarschule sind. Andere Variablen auf Individualebene hatten hingegen nur bei einzelnen Messzeitpunkten oder nur in einer der beiden Untersuchungsstichproben einen signifikanten Einfluss auf die Mathematikleistung:

- Einschränkungen im Deutschverstehen (beim Post-Test und beim Follow-up 2 in der Untersuchungsstichprobe Stichprobe^{IQstreng}),
- Alter (beim Follow-up 1 in beiden Untersuchungsstichproben).

Bei keinem Messzeitpunkt hat sich hingegen gezeigt, dass der SES ein signifikanter Prädiktor für die Mathematikleistung auf Individualebene ist.

Auf *Klassenebene* ergaben die Analysen für die untersuchten Prädiktoren, dass einzelne in dieser Arbeit berücksichtigte Variablen zwar einen signifikanten Einfluss auf die Mathematikleistung haben, jedoch nur bei bestimmten Messzeitpunkten und teilweise nur in einer der beiden Untersuchungsstichproben:

- IQ-Mittelwert der Klassen (beim Post-Test und beim Follow-up 1 in beiden Untersuchungsstichproben),
- Vorwissen der Klasse (beim Post-Test in der Stichprobe^{IQstreng}; kein signifikanter Einfluss, wenn IQ-Mittelwert der Klasse ebenfalls berücksichtigt wird),
- Anteil Knaben in den Klassen (beim Follow-up 1 in beiden Untersuchungsstichproben),
- Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen (beim Follow-up 1 in beiden Untersuchungsstichproben).

Einzelne aggregierte Klassenvariablen, wie sie in der vorliegenden Arbeit für die Analysen verwendet wurden, können somit einen signifikanten Einfluss auf die individuelle Mathematikleistung in der dritten und der vierten Klasse der Primarschule haben. Eine theoretische Einordnung der Ergebnisse sowie mögliche Gründe für die Ergebnisse werden im Folgenden jeweils separat für die einzelnen Variablen aufgezeigt.

7.2.1 Mathematisches Vorwissen

Für das mathematische Vorwissen zeigt sich auf *Individualebene* für beide Untersuchungsstichproben ein deutlich positiver Einfluss auf die Mathematikleistung, wonach Kinder mit höheren Werten im Vorwissen auch bei allen drei später erfolgten Messzeitpunkten über höhere Mathematikleistungen verfügen (vgl. Kap. 6.2). Dies steht im Einklang mit empirischen Forschungserkenntnissen, wonach das Vorwissen ein zentraler Prädiktor für spätere Mathematikleistungen ist (vgl. Kap. 2.3.2), und stimmt auch mit entwicklungstheoretischen Überlegungen überein (vgl. Kap. 2.2), wonach sich bereits erworbene mathematische Kompetenzen positiv auf die weitere mathematische Entwicklung auswirken. Vergleicht man die standardisierten Regressionskoeffizienten für das mathematische Vorwissen der Stichprobe^{alle} für die verschiedenen Messzeitpunkte, so zeigt sich, dass der durchschnittliche Einfluss des mathematischen Vorwissens erfasst am Ende der zweiten Klasse für alle drei nachfolgenden Messzeitpunkte ähnlich hoch ausfällt (Post-Test: $\beta = 0.57$; Follow-up 1: $\beta = 0.60$; Follow-up 2: $\beta = 0.57$) und vergleichbar ist mit den Werten für die Stichprobe^{IQstreng} (Post-Test:[30] $\beta = 0.55$; Follow-up 1: $\beta = 0.60$; Follow-up 2: $\beta = 0.58$). Der Einfluss des individuellen Vorwissens in Mathematik kann somit für die dritte und die vierte Klasse als stabil eingeschätzt werden. Die Werte der standardisierten Regressionskoeffizienten für das Vorwissen aus der vorliegenden Untersuchung bewegen sich zudem in einer ähnlichen Grössenordnung wie der Wert, der in der Studie von Moser Opitz et al. (2017, S. 7) für den Post-Test angegeben wird ($\beta = 0.61$; vgl. Kap. 3.1.4).

Bezüglich des Einflusses des mathematischen Vorwissens muss jedoch berücksichtigt werden, dass der Zusammenhang zwischen mathematischem Vorwissen und Mathematikleistung beim Follow-up 2 in der Stichprobe^{alle} zwischen den Klassen variiert. So hat das mathematische Vorwissen in dieser Untersuchungsstichprobe in gewissen Klassen einen knapp viermal so hohen Einfluss auf die individuelle Mathematikleistung wie in Klassen mit dem geringsten Einfluss des Vorwissens. Dies könnte dadurch bedingt sein, dass der Lern- und Wissenszuwachs in einzelnen Klassen aufgrund qualitativ unterschiedlich guten Unterrichts bis Ende vierter Klasse grösser war als in anderen Klassen (vgl. Waldis, Grob, Pauli & Reusser, 2010), sodass sich daher für verschiedene Klassen differenzielle Ergebnisse beim Follow-up 2 ergeben haben.

Auf *Klassenebene* zeigt sich in der vorliegenden Arbeit, dass das Vorwissen der Klasse ebenfalls ein signifikanter Prädiktor für die individuelle Mathematikleistung ist, dies jedoch nur auf den Post-Test und nur auf die Stichprobe^{IQstreng} zutrifft, wenn nicht gleichzeitig der IQ-Mittelwert für die Klasse mitmodelliert wird (vgl. Mehrebenenmodelle in Kap. 6.2.2). Das weist darauf hin, dass das Vorwissen und der IQ-Mittelwert der Klasse einen gemeinsamen Anteil der Varianz in den individuellen Mathematikleistungen erklären. Zudem wiederspiegeln sich darin entsprechende Effekte auf Individualebene in abgeschwächter Form (vgl. hierzu auch Kap. 7.2.2). Vergleicht man beim Post-Test den Wert für den standardisierten Regressionskoeffizienten des Vorwissens auf Individualebene ($\beta = 0.55$) mit demjenigen auf Klassenebene ($\beta = 0.09$), so ist Letzterer rund sechsmal kleiner. Daraus kann geschlossen werden, dass dem individuellen Vorwissen eine grössere Erklärungskraft für die Mathematikleistung zukommt als dem Vorwissen der Klasse. Zwar haben vereinzelt auch andere Untersuchungen einen Einfluss

[30] Beim Post-Test wurde der standardisierte Regressionskoeffizient auf Basis des Modells 4b berechnet.

des Vorwissens der Klasse auf die individuellen Leistungen finden können (z.B. Nikolova, 2011; Zurbriggen, 2016), diesbezüglich kann jedoch nicht von einem gut belegten Forschungsstand ausgegangen werden.

7.2.2 Intelligenz

Auch die Intelligenz hat sich auf *Individualebene* unter zusätzlicher Berücksichtigung des Vorwissens in Mathematik als signifikanter Prädiktor für die Mathematikleistung in der dritten und zu Beginn der vierten Klasse der Primarschule herausgestellt. Demnach erreichen Schülerinnen und Schüler mit höheren Intelligenzwerten auch höhere Mathematikleistungen. Aufgrund der Erkenntnisse aus Kapitel 2.3.2 ist das Ergebnis, wonach die Intelligenz unter zusätzlicher Berücksichtigung des mathematischen Vorwissens ein signifikanter Prädiktor für die Mathematikleistung ist, jedoch eher als erwartungswidrig einzuschätzen. Allerdings liegt hierzu auch eine Studie vor, die für rechenschwache Schülerinnen und Schüler ebenfalls einen Einfluss der Intelligenz auf die Mathematikleistung unter zusätzlicher Berücksichtigung des Vorwissens belegen konnte (Moser Opitz et al., 2017). Da in der vorliegenden Arbeit eine leistungsheterogene Stichprobe untersucht wurde, wäre eine mögliche Hypothese, dass bei Schülerinnen und Schülern mit wenig mathematischem Vorwissen der Intelligenz eine grössere Erklärungskraft zukommt, als wenn leistungsstärkere Schülerinnen und Schüler untersucht werden.

Vergleicht man die standardisierten Regressionskoeffizienten des IQs am Beispiel der Stichprobe^{alle} (Post-Test: $\beta = 0.12$; Follow-up 1: $\beta = 0.08$; Follow-up 2: $\beta = 0.14$) mit den standardisierten Regressionskoeffizienten des Vorwissens (Post-Test: $\beta = 0.57$; Follow-up 1: $\beta = 0.60$; Follow-up 2: $\beta = 0.57$), kann festgestellt werden, dass der durchschnittliche Einfluss des IQs auf die Mathematikleistung um einiges geringer ausfällt als jener des mathematischen Vorwissens. Dies stimmt mit Erkenntnissen aus der Forschung überein (vgl. Kap. 2.3.2). Interessant ist diesbezüglich auch ein Vergleich mit der Studie von Moser Opitz et al. (2017, 7f), in der nur rechenschwache Schülerinnen und Schüler untersucht wurden. Dort fallen die standardisierten Regressionskoeffizienten für den IQ doppelt so hoch aus wie in der vorliegenden Untersuchung (Post-Test: $\beta = 0.20$; Follow-up: $\beta = 0.24$; Moser Opitz et al., 2017, 7f). Dies untermauert die oben aufgestellte Hypothese, wonach bei Kindern mit Rechenschwäche die Intelligenz einen grösseren Einfluss auf die mathematische Leistungsentwicklung zu haben scheint, als wenn Schülerinnen und Schüler aus einem heterogenen Leistungsspektrum untersucht werden (vgl. hierzu auch Helmke & Weinert, 1997). Es kann daher angenommen werden, dass bei rechenschwachen Kindern, die über weniger mathematisches Vorwissen verfügen, der Einfluss der

Intelligenz auf die mathematische Leistungsentwicklung im Vergleich zum mathematischen Vorwissen höher ausfällt, als wenn man alle Schülerinnen und Schüler einer Klasse untersucht.

Der IQ hat auch auf *Klassenebene* in Form des IQ-Mittelwerts der Klasse einen signifikanten Einfluss auf die Mathematikleistung, wobei dies jedoch in beiden Untersuchungsstichproben nur auf den Post-Test und den Follow-up 1 zutrifft. Demnach kann für den IQ ein sogenannter Kompositionseffekt (vgl. hierzu z.B. Dumont, Neumann, Maaz & Trautwein, 2013) angenommen werden, wonach sich die Klassenzusammensetzung bezüglich des Intelligenzniveaus zusätzlich zum individuellen Intelligenzniveau ebenfalls auf die individuelle Mathematikleistung auswirkt. Dabei zeigt sich in der vorliegenden Untersuchung, dass ein hoher IQ-Mittelwert der Klasse die individuelle Mathematikleistung beim Post-Test und beim Follow-up 1 positiv beeinflusst und Kinder somit von einem hohen Intelligenzniveau der Klasse profitieren. Dies steht im Einklang mit Forschungsergebnissen aus anderen Untersuchungen (z.B. Lehmann, 2006; Opdenakker, van Damme, de Fraine, van Landeghem & Onghena, 2002; Tiedemann & Billmann-Mahecha, 2004).

Vergleicht man beispielhaft für die Stichprobe^{alle} die standardisierten Regressionskoeffizienten für den IQ-Mittelwert der Klasse (Post-Test: $\beta = 0.10$; Follow-up 1: $\beta = 0.09$) mit den standardisierten Regressionskoeffizienten für den IQ als Individualprädiktor (Post-Test: $\beta = 0.12$; Follow-up 1: $\beta = 0.08$), kann festgestellt werden, dass das Intelligenzniveau der Klasse die Mathematikleistung in der vorliegenden Untersuchung ähnlich stark beeinflusst wie die individuelle Intelligenz. Im Vergleich zum Einfluss des individuellen Vorwissens ist der kumulierte Einfluss der individuellen und der klassenbezogenen Intelligenz jedoch immer noch ca. dreimal geringer. Demnach kommt dem individuellen Vorwissen eine höhere Erklärungskraft zu als dem kumulierten Effekt des individuellen und klassenbezogenen Intelligenzniveaus.

7.2.3 Geschlecht

Beim Geschlecht ergeben sich auf *Individualebene* für alle drei Messzeitpunkte im Durchschnitt höhere Leistungen für Knaben, sodass das Geschlecht sowohl in der dritten Klasse als auch in der vierten Klasse ein signifikanter Prädiktor für die Mathematikleistung ist. Dies entspricht der Datenlage für die Schweiz, steht jedoch nicht im Einklang mit Ergebnissen aus internationalen Studien, wonach über verschiedene kulturelle Kontexte hinweg keine Geschlechterunterschiede festzustellen sind (vgl. Kap. 2.3.4). Wie Else-Quest et al. (2010) aufzeigen konnte, haben Faktoren wie die Gleichstellung der Geschlechter in einem Land einen Einfluss darauf, ob Geschlechterunterschiede bezüglich Mathematikleistung in den einzelnen Ländern gefunden werden. Gemäss den Daten, die in diesem Artikel für die Schweiz präsentiert werden, können die Ergebnisse aus der vorliegenden Untersuchung daher als erwartungskonform eingeschätzt werden. Inwiefern diese Ergebnisse allenfalls auch durch unterschiedlich ausgeprägte Einstellungen im Fach Mathematik (z.B. Angst vor Mathematik, Selbstvertrauen in Mathematik) bei Knaben und Mädchen beeinflusst sind (vgl. Kap. 2.3.6), kann im Rahmen der vorliegenden Untersuchung jedoch nicht geklärt werden, da entsprechende Daten nicht erhoben wurden.[31]

Auch auf *Klassenebene* zeigt sich für das Geschlecht ein signifikanter Einfluss auf die individuelle Mathematikleistung. Dies konnte für beide Untersuchungsstichproben festgestellt werden, trifft aber nur für den Follow-up 1 zu (vgl. Kap. 6.2). Dabei zeigt sich, dass sich ein über dem Stichprobendurchschnitt liegender Anteil an Knaben in den Klassen positiv auf die individuelle Mathematikleistung ausgewirkt hat. Dieser Effekt ist somit konform mit dem Effekt auf Individualebene. Im Vergleich zur Individualebene muss jedoch davon ausgegangen werden, dass der Geschlechteranteil in den Klassen ein weniger eindeutiger Prädiktor für die individuelle Mathematikleistung ist als das individuelle Geschlecht, da sich dieser Zusammenhang nur bei einem Messzeitpunkt gezeigt hat.

7.2.4 Förderbedarf in Mathematik

Ob ein Kind einen Förderbedarf in Mathematik hat oder nicht, hat sowohl in der dritten Klasse als auch am Ende der vierten Klasse der Primarschule einen signifikanten Einfluss auf die Mathematikleistung. Es konnte demnach bestätigt werden, dass Kinder mit einem Förderbedarf in Mathematik erwartungsgemäss über niedrigere Mathematikleistungen verfügen (vgl. Kap. 2.1.3), wobei dieser Unterschied im Durchschnitt zwischen 3.57 und 4.20 Punkte im jeweiligen Mathematiktest beträgt (vgl. Kap. 6.2). Allerdings zeigt sich beim Follow-up 2 in beiden Untersuchungsstichproben, dass der Zusammenhang zwischen Förderbedarf in Mathematik und Mathematikleistung am Ende der vierten Klasse nicht in allen Klassen gleich ausfällt. Auf Basis des 95%-Konfidenzintervalls lässt sich schliessen, dass sich in gewissen Klassen ein beim Pre-Test erfasster Förderbedarf sogar positiv auf die Mathematikleistung von Schülerinnen und Schülern mit Förderbedarf beim Follow-up 2 ausgewirkt hat. Ein Teil dieser Varianz kann durch den Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen erklärt werden (*Cross-Level-Interaktion*; vgl. Kap. 6.2). Demnach zeigt sich für Klassen mit einem hohen Anteil an Kindern mit Einschränkungen im Deutschverstehen, dass ein Förderbedarf mit

[31] Spannend wäre an dieser Stelle ein Vergleich der standardisierten Regressionskoeffizienten für das Geschlecht mit denjenigen für das mathematische Vorwissen und den IQ. Da bei Dummy-codierten Variablen aufgrund der Metrik jedoch standardisierte Regressionskoeffizienten nicht sinnvoll interpretiert werden können, wird auf diesen Vergleich verzichtet (vgl. Kap. 5.6.2).

niedrigeren Mathematikleistungen einhergeht als in Klassen mit einem unterdurchschnittlichen Anteil an Kindern mit Einschränkungen im Deutschverstehen. Da individuelle sprachliche Kompetenzen die Mathematikleistungen negativ beeinflussen können (vgl. Kap. 2.3.3), ist nicht gänzlich überraschend, dass ein Teil der Varianz zwischen den Klassen mit Bezug auf den Zusammenhang zwischen Förderbedarf und Mathematikleistungen durch das sprachliche Niveau der Klasse in Form eines hohen Anteils an Schülerinnen und Schülern mit Einschränkungen im Deutschverstehen erklärt werden kann.

Da der Anteil Schülerinnen und Schüler mit Einschränkungen im Deutschverstehen jedoch nur einen Teil der Varianz aufklärt, kommen auch weitere Erklärungsmöglichkeiten infrage. So könnte die Feststellung, dass in gewissen Klassen Schülerinnen und Schüler mit einem Förderbedarf beim Post-Test ihre Mathematikleistung in der Zeit bis zum Follow-up 2 bedeutsam steigern konnten, dadurch bedingt sein, dass Kinder mit Förderbedarf in Mathematik aufgrund dieses Status zusätzlich eine wirkungsvolle Unterstützungsmassnahmen im Fach Mathematik erhalten haben und sich dies zwei Jahre später positiv auf die Mathematikleistung ausgewirkt hat (Moser Opitz, 2008a). Dass in gewissen Klassen das Gegenteil der Fall ist (d.h., dass in diesen ein Förderbedarf mit sehr niedrigen Leistungen beim Follow-up 2 einhergeht), wäre in diesem Sinne dadurch zu erklären, dass in diesen Klassen Schülerinnen und Schüler mit Förderbedarf keine adäquate Unterstützung im Fach Mathematik erhalten haben und sich die Leistungsrückstände für diese Schülerinnen und Schüler entsprechend noch vergrössert haben. Aufgrund der föderalistischen Struktur des schweizerischen Schulsystems kann sich neben der jeweiligen Qualität der Unterstützungsmassnahmen auch das Ausmass der Unterstützung je nach Schulstandort teilweise stark unterscheiden (Moser Opitz, 2008a). Es kann somit als plausibel angenommen werden, dass nebst unterschiedlicher Unterstützungsqualität auch abweichende strukturelle Rahmenbedingungen hinsichtlich der Quantität der Unterstützung zu den oben beschriebenen Unterschieden beigetragen haben. Eine weitere mögliche Erklärung lautet, dass es sich bei einem Teil der Kinder mit Förderbedarf in Mathematik um Kinder handelt, die nur aufgrund zeitlich begrenzter Schwierigkeiten im Fach Mathematik einen solchen Status erhalten haben (vgl. Mazzocco & Räsänen, 2013; Vukovic & Siegel, 2010), sodass sich für diese Kinder im Vergleich zum Ende der zweiten Klasse am Ende der vierten Klasse kein Leistungsrückstand mehr zeigt.

7.2.5 Einschränkungen im Deutschverstehen

Auf *Individualebene* erweisen sich Einschränkungen im Deutschverstehen nur für die Stichprobe^{IQstreng} und nur beim Post-Test und beim Follow-up 2 als signifikanter Prädiktor für die Mathematikleistung. Dass dies nicht bei allen Messzeitpunkten und nur in einer Untersuchungsstichprobe der Fall ist, steht nicht im Einklang mit Ergebnissen, wie sie in Kapitel 2.3.3 geschildert wurden. Dass in der vorliegenden Untersuchung Einschränkungen im Deutschverstehen sich nicht durchgängig als signifikanter Prädiktor für die Mathematikleistung erwiesen haben, könnte durch die Operationalisierung beeinflusst sein. So wurde hier als Indikator für Sprachkompetenzen eine Einschätzung der Lehrpersonen zum Deutschverständnis gewählt, in anderen Studien wurden jedoch die Sprachkompetenz bzw. die Lesekompetenz in Form von standardisierten Leistungstests erfasst oder indirekt mithilfe der Familiensprache erhoben (vgl. Kap. 2.3.3 und Kap. 5.4.4).

Da sich für die Stichprobe$^{\text{alle}}$ und die Stichprobe$^{\text{IQstreng}}$ unterschiedliche Ergebnisse gezeigt haben, kann vermutet werden, dass die Zusammensetzung der Schülerinnen und Schüler in diesen beiden Stichproben eine Rolle gespielt hat. In der Stichprobe$^{\text{alle}}$, für die kein signifikanter Einfluss von Einschränkungen im Deutschverstehen auf die Mathematikleistung festgestellt werden konnte, weisen die untersuchten Kinder im Durchschnitt weniger mathematisches Vorwissen und einen niedrigeren IQ auf (vgl. Kap. 5.5). Da beide Prädiktoren einen signifikanten Einfluss auf die Mathematikleistung haben, wäre es denkbar, dass bei Schülerinnen und Schülern mit ungünstigeren kognitiven Lernvoraussetzungen sich Einschränkungen im Deutschverstehen nicht zusätzlich negativ auf die Mathematikleistungen auswirken. Untersucht man hingegen Schülerinnen und Schüler mit höheren kognitiven Lernvoraussetzungen, wie dies in der Stichprobe$^{\text{IQstreng}}$ der Fall ist, können Einschränkungen im Deutschverstehen einen zusätzlichen Beitrag zur Erklärung individueller Unterschiede in der Mathematikleistung liefern. Das würde heissen, dass ein geringes mathematisches Vorwissen für die Mathematikleistung bedeutsamer ist als Einschränkungen im Deutschverstehen, sich hingegen bei Kindern mit vergleichsweise höherem mathematischem Vorwissen unterschiedliche Deutschkenntnisse durchaus auf die Mathematikleistung auswirken können.

Auch auf *Klassenebene* zeigen sich Einschränkungen im Deutschverstehen teilweise als bedeutsam für die individuelle Mathematikleistung. So hat der Anteil an Kindern mit Einschränkungen im Deutschverstehen beim Follow-up 1 in beiden Untersuchungsstichproben einen signifikanten Einfluss auf die Mathematikleistung. Dabei zeigt sich, dass sich ein hoher Anteil an Kindern mit Einschränkungen im Deutschverstehen negativ auf die individuelle Mathematikleistung auswirkt. Allerdings handelt es sich hier nicht um einen eigentlichen Kompositionseffekt, da im Kapitel 6.3 unter zusätzlicher Berücksichtigung von Einschränkungen im Deutschverstehen auf Individualebene kein Effekt für den Anteil an Kindern mit Einschränkungen im Deutschverstehen in den Klassen mehr festgestellt werden konnte.[32] Von einem Kompositionseffekt wird nur dann gesprochen, wenn ein Merkmal sowohl auf Individual- als auch auf Klassenebene einen Einfluss auf die abhängige Variable hat (vgl. Dumont et al., 2013; Stanat, 2006).

An dieser Stelle soll auch nochmals darauf hingewiesen werden, dass beim Follow-up 2 in beiden Untersuchungsstichproben festgestellt wurde, dass ein hoher Anteil an Kindern mit Einschränkungen im Deutschverstehen in den Klassen einen Teil der Varianz zwischen den Klassen mit Bezug auf den Zusammenhang zwischen Förderbedarf und Mathematikleistung erklären konnte (vgl. die Diskussion zum Förderbedarf in Mathematik).

7.2.6 Alter

Das Alter hat sich auf *Individualebene* nur beim Follow-up 1 am Ende der dritten Klasse als signifikanter Prädiktor für die Mathematikleistung in beiden Untersuchungsstichproben erwiesen, wobei ältere Schülerinnen und Schüler durchschnittlich über niedrigere Mathematikleistungen verfügen. Dies wirkt im ersten Moment überraschend, da grundsätzlich von einer Zunahme der mathematischen Leistungen über die Zeit ausgegangen werden kann (z.B. Ehlert, Fritz, Arndt & Leutner, 2013). Allerdings beziehen sich entsprechende Untersuchungen auf Schülerinnen und Schüler aus unterschiedlichen Klassenstufen, sodass das Alter in diesen Fällen ein Indikator für zusätzlich behandelten Lernstoff und somit für zusätzliche Lerngelegenheiten ist (Baumert, Brunner, Lüdtke & Trautwein, 2007). Daher kommen auch Helmke und Weinert (1997, 102f) zu folgendem Schluss:

> Obwohl es während der Kindheit zwischen Lebensalter und fast allen kognitiven Leistungen positive, numerisch durchwegs substanzielle […] korrelative Zusammenhänge gibt, ist und bleibt das chronologische Alter eine inhaltsleere psychologische Variable, die nur durch ihren Hinweischarakter auf Reifungsvorgänge, Lernprozesse und Bildungseinflüsse praktische Bedeutung gewinnt.

Zwar kann auch für die vorliegende Untersuchung angenommen werden, dass ein grosser Teil der älteren Schülerinnen und Schüler aufgrund von Repetitionen oder dem Besuch von Einführungsklassen im Vergleich zu jüngeren Schülerinnen und Schülern zusätzliche Lerngelegenheiten erhalten haben. Allerdings hat sich das für diese Schülerinnen und Schüler nicht in Form eines Leistungsvorsprungs ausgewirkt. Dies kann jedoch als plausibel eingeschätzt werden, da vermutet werden kann, dass diese Schülerinnen und Schüler aufgrund von Lernschwierigkeiten

[32] Zusätzlich wurden in diesen Modellen auch der SES und die Dummy-Variablen für die Stichprobengruppenzugehörigkeit berücksichtigt.

eine Klasse repetieren mussten oder später eingeschult wurden und somit mit erhöhter Wahrscheinlichkeit nach wie vor über niedrigere schulische Leistungen auch im Fach Mathematik verfügen. So belegen diverse Studien, dass sich Repetitionen längerfristig nicht positiv auf die schulischen Leistungen u.a. auch im Fach Mathematik auswirken (Bless, Schüpbach & Bonvin, 2004; Hess, 2010; Jimerson, 2001). Ein im Vergleich zur restlichen Stichprobengruppe höheres Alter wäre im Sinne von Helmke und Weinert (1997) demnach ein Indikator für eine Repetition oder eine spätere Einschulung und dies wiederum ein Indikator für längerfristige schulische Schwierigkeiten. Auch in einer anderen Studie wurde ein überdurchschnittlich hohes Alter gewisser Kinder in gleicher Art und Weise interpretiert, wobei sich das Alter dort nicht als signifikanter Prädiktor für die Mathematikleistung erwiesen hat (Tiedemann & Billmann-Mahecha, 2004).

7.2.7 Sozioökonomischer Status (SES)

Bei keinem Messzeitpunkt und in keiner der beiden Untersuchungsstichproben hat sich der SES als signifikanter Prädiktor für die Mathematikleistung erwiesen. Dies steht nicht im Einklang mit Ergebnissen aus anderen Untersuchungen (vgl. Kap. 2.3.5). Begründet wird der Einfluss des SES auf schulische Leistungen damit, dass sich je nach SES der Familie z.B. familiäre Lerngelegenheiten oder elterliche Bildungsaspirationen unterscheiden können (vgl. Jordan & Levine, 2009). Allerdings muss für die vorliegende Untersuchung berücksichtigt werden, dass für die Analysen Variablen eingesetzt wurden, für die angenommen werden kann, dass sie einen (indirekten) Indikator für den SES darstellen können (2.3.5). Eine Hypothese könnte daher lauten, dass der SES keine zusätzliche Erklärungskraft aufweist, wenn indirekte Indikatoren für den SES wie Einschränkungen im Deutschverstehen mitmodelliert werden. In diese Richtung weisen auch die Ergebnisse für die Stichprobe^{IQstreng} aus den Modellen, die im Kapitel 6.3.2 für die Überprüfung der Wirksamkeit der Förderung präsentiert wurden und in denen der SES als Kontrollvariable eingefügt wurde. So haben Einschränkungen im Deutschverstehen beim Post-Test und beim Follow-up 2 nur noch einen tendenziellen Einfluss auf die Mathematikleistung, wenn der SES mitmodelliert wird, Gleiches trifft beim Follow-up 1 auf den Anteil Kinder mit Einschränkungen im Deutschverstehen in den Klassen zu. Es kann daher vermutet werden, dass der SES und Einschränkungen im Deutschverstehen einen gemeinsamen Erklärungsanteil aufweisen.

Eine andere Erklärungsmöglichkeit wäre, dass die Operationalisierung des SES über die Bücheraufgabe von Paulus (2009) zu den Ergebnissen beigetragen hat. So wird in einigen berichteten Studien eine komplexere Erfassung des SES, z.B. über das Einkommen oder den Bildungsabschluss der Eltern, vorgenommen (vgl. Kap. 2.3.5). Im Gegensatz dazu muss eine Einschätzung des familiären Bücherbestandes durch die Schülerinnen und Schüler als weniger reliabel eingeschätzt werden. Für zukünftige Untersuchungen ist diesbezüglich auch zu beachten, dass der familiäre Bücherbestand in Form von Büchern in einem Bücherregal angesichts der zunehmenden Digitalisierung in absehbarer Zeit nicht mehr zu einer zuverlässigen Einschätzung des kulturellen Kapitals im Sinne Bourdieus (1983) führen könnte.

7.3 Grenzen der Untersuchung

Bevor ein abschliessendes Fazit gezogen wird, werden an dieser Stelle noch Grenzen der vorliegenden Untersuchung dargestellt. So muss mit Blick auf die Frage nach signifikanten Prädiktoren für die Mathematikleistung festgehalten werden, dass nicht alle potenziellen Einflussfaktoren berücksichtigt werden konnten. Zu nennen ist diesbezüglich insbesondere das Arbeitsgedächtnis. Einerseits kann angenommen werden, dass eine eingeschränkte Arbeitsgedächtniskapazität auf Individualebene die Mathematikleistung negativ beeinflussen kann (vgl. Kap. 2.3.1). Andererseits hat sich in der Studie von Powell, Cirino und Malone (2017) gezeigt, dass rechenschwache Kinder mit hoher Arbeitsgedächtniskapazität stärker von einer Förderung im Fach Mathematik profitieren als rechenschwache Kinder mit niedriger Arbeitsgedächtniskapazität. Da für die Erfassung des Arbeitsgedächtnisses jedoch keine Gruppentests existieren, musste aus ökonomischen Gründen auf die Erhebung dieses Einflussfaktors verzichtet werden.

Ebenfalls nicht erfasst wurden emotionale und motivationale Aspekte (z.B. Angst vor Mathematik). Es kann angenommen werden, dass diese ebenfalls einen Einfluss auf die Leistungen im Fach Mathematik haben (vgl. Kap. 2.3.6). Angesichts der fehlenden Interventionseffekte der unterrichtsintegrierten Förderung mit Blick auf die Mathematikleistung von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen wäre es zudem auch interessant emotionale bzw. motivationale Faktoren als abhängige Variable zu erfassen. So haben Lehrpersonen der Interventionsgruppen berichtet, dass sie im emotionalen und motivationalen Bereich Veränderungen bei Kindern mit mathematischen Schwierigkeiten wahrgenommen haben. Es ist daher möglich, dass die unterrichtsintegrierte Förderung für diese Schülerinnen und Schüler zu wenig intensiv bzw. spezifisch war, um sich positiv auf die Mathematikleistung in einem standardisierten Test auszuwirken, dass aber emotional und motivational eine Verbesserung der Situation erreicht werden konnte, was sich längerfristig wiederum positiv auf die Leistungsentwicklung auswirken sollte (vgl. Kap. 2.3.6).

Bezüglich der Untersuchung der Effektivität der Förderung muss zudem beachtet werden, dass auf Klassenebene keine Indikatoren für die Qualität des Unterrichts berücksichtigt wurden.

Zwar wurde im Rahmen des Gesamtprojektes, auf das sich die vorliegende Untersuchung bezieht, die Adaptivität der Lernbegleitung durch die Lehrperson mithilfe von Videoanalysen und einem hochinferenten Rating analysiert (Pfister, 2016). Da jedoch für die Kontrollklassen keine Aufnahmen vorliegen, konnte dieser Qualitätsaspekt in den Mehrebenenanalysen nicht untersucht werden. Weiter wurden auch Merkmale der Lehrpersonen (z.B. professionelle Kompetenzen, Berufserfahrung, Einstellungen zum Fach) nicht für die Analysen berücksichtigt (vgl. Kap. 5.2). Im Rahmen des Gesamtprojektes wurde das Wissen der Lehrpersonen zum Thema Rechenschwäche zwar erfasst (z.B. typische Fehler, Ursachen, Fördermöglichkeiten). Allerdings wurden diese Angaben anonym erhoben, um eine möglichst hohe Beteiligung am Fragebogen zu erreichen, sodass diese Daten auch nicht für die Mehrebenenanalysen verwendet werden konnten (vgl. Kap. 5.2). Ein weiterer stichprobenspezifischer Unterschied, der allenfalls zur Erklärung der abweichenden Ergebnisse für die Gruppe^{Begl} und die Gruppe^{Mat} infrage kommen könnte, ist die Anzahl Mischklassen, die in der Gruppe^{Begl} um einiges höher ist als in den beiden anderen Stichprobengruppen. Dieser Aspekt wurde in den vorliegenden Analysen jedoch nicht berücksichtigt, sodass hierzu keine Aussagen gemacht werden können.

Ebenfalls einschränkend zu erwähnen ist, dass bei der Analyse der Effektivität der Förderung mit Blick auf Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen die hierarchische Datenstruktur nicht berücksichtigt werden konnte, da die Fallzahlen pro Klasse dafür zu niedrig sind. So wäre es interessant zu untersuchen, ob bei diesen Schülerinnen und Schülern die gleichen Prädiktoren auf Klassenebene einen Einfluss auf die individuellen Mathematikleistungen haben oder ob sich hier Unterschiede zu den Ergebnissen zeigen, bei denen Schülerinnen und Schüler des gesamten Leistungsspektrums berücksichtigt wurden. Ebenso konnte für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen nicht analysiert werden, ob sich für einzelne Klassen differenzielle Ergebnisse für die Stärke und die Richtung der Zusammenhänge zwischen den Prädiktorvariablen und der abhängigen Variable ergeben.

7.4 Fazit

Aufgrund der Ergebnisse der in dieser Arbeit untersuchten mathematischen Interventionsstudie wird für zukünftige Studien, in denen eine mathematische Förderung für Schülerinnen und Schüler mit unterdurchschnittlichen Mathematikleistungen in Regelklassen umgesetzt werden soll, empfohlen, dass inhaltliche und (fach-)didaktische Aspekte, die sich auch in anderen Interventionsstudien als effektiv erwiesen haben, beibehalten werden sollten. Dazu gehört eine verstehensorientierte Fokussierung auf zentrale mathematische Kompetenzen und Inhalte, eine starke Strukturierung der Förderung und der Einsatz von produktiven Übungsformaten sowie von geeigneten Veranschaulichungen und Arbeitsmitteln. Für zukünftige Interventionsstudien ist hingegen zu überlegen, in welches Verhältnis gemeinsame Lernsituationen und individuelle Förderung im Rahmen einer unterrichtsintegrierten Mathematikförderung zu bringen sind. So sind zusätzliche Individualisierungs- und Differenzierungsmassnahmen ebenso zu prüfen, wie Massnahmen, die eine optimalere Passung zwischen individuellen Lernvoraussetzungen und mathematischen Fördermassnahmen ermöglichen. Auch der punktuelle Einsatz von Klein- oder Einzelgruppensequenzen sollte für zukünftige Untersuchungen in Betracht gezogen werden, damit auf die individuellen Bedürfnisse von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen trotz grundsätzlich unterrichtsintegriertem Fördersetting stärker eingegangen werden kann. Dabei sollte auch der Einbezug Schulischer Heilpädagoginnen und Heilpädagogen bzw. für Untersuchungen in Deutschland die Mitarbeit von Sonderschullehrkräften für die Umsetzung der Förderung erwägt werden. Dabei ist deren Einsatz sowohl innerhalb des Klassenzimmers als auch punktuell ausserhalb des Klassenzimmers denkbar. In diesem Zusammenhang wäre auch interessant zu untersuchen, ob diese Lehrpersonen über zusätzliches Wissen bzw. zusätzliche Kompetenzen für die Förderung und die individuelle Lernbegleitung von Schülerinnen und Schülern mit unterdurchschnittlichen Mathematikleistungen verfügen.

Die Ergebnisse der vorliegenden Arbeit bieten hinsichtlich individueller und klassenbezogener Einflussfaktoren für mathematische Leistungen zudem spannende Erkenntnisse. So hat sich gezeigt, dass das Vorwissen in gewissen Klassen einen viel grösseren Einfluss auf die mathematischen Leistungen hat als in anderen Klassen, wobei Unterschiede hinsichtlich der Unterrichtsqualität als mögliche Erklärung dafür in Frage kommen. Ebenso konnte bestätigt werden, dass unter Berücksichtigung eines breiteren Leistungsspektrums die allgemeine Intelligenz die mathematischen Leistungen stärker beeinflusst, als wenn dies in Regeklassen ohne Einbezug von Lernenden mit unterdurchschnittlichen Mathematikleistungen untersucht wird. Des Weiteren zeigt die vorliegende Arbeit, dass auch für zukünftige Interventionsstudien auf Klassenebene aggregierte Individualmerkmale in die Analysen einbezogen werden sollten, da diese die individuellen Mathematikleistungen beeinflussen können. Aufgrund der vorliegenden Ergebnisse muss angenommen werden, dass es sich dabei sowohl auf Individual- als auch auf Klassenebene um komplexe Zusammenhänge zwischen verschiedenen Merkmalen handelt. So zeigte sich in der vorliegenden Studie ein Einfluss des Anteils an Lernenden mit Einschränkungen im Deutschverstehen in den Klassen auf den Zusammenhang zwischen Förderbedarf in Mathematik und den individuellen Mathematikleistungen. Mit Blick auf Individual- und Klassenvariablen, die die mathematische Leistung beeinflussen, sollte zudem geprüft werden, wie diese Faktoren bei der Gestaltung der mathematischen Förderung noch besser berücksichtigt werden können. So ist hinsichtlich der didaktischen und mathematikdidaktischen Gestaltung einer verstehensorientierten Förderung, in der viel Wert auf die Versprachlichung sowie die gemeinsame Reflexion mathematischer Rechenwege und zentraler mathematischer Konzepte gelegt wird, zu überlegen, wie Schülerinnen und Schüler, die die Schulsprache nicht oder nur eingeschränkt verstehen, von entsprechenden Sequenzen besser profitieren können. Eine stärkere Sensibilisierung der Lehrpersonen für solche erschwerenden Bedingungen könnte hierfür bereits ein erster hilfreicher Ansatz sein.

Die vorliegende Arbeit bietet für zukünftige Forschungsprojekte darüber hinaus weitere wichtige forschungsmethodische Hinweise. So wird eine genauere Erfassung der Qualität, mit der die Förderung in den einzelnen Klassen umgesetzt wird, angesichts der hier gefundenen widersprüchlichen Ergebnisse für die beiden Interventionsgruppen als zentral erachtet. Unerlässlich ist dabei, dass auch die Kontrollgruppe einbezogen wird und bei einer längeren Interventionszeit nach Möglichkeit mehrere Erhebungen der Umsetzungsqualität stattfinden. Die Ergebnisse der vorliegenden Arbeit belegen zudem, dass es von grosser Wichtigkeit ist Merkmale auf Individual- und Klassenebene zu erheben, die einen Einfluss auf die mathematische Leistung haben. Für zukünftige Studien sollte dabei auch geprüft werden, welche zusätzlichen Variablen direkt auf Klassenebene erfasst werden können (Faktoren für die Unterrichtsqualität, Wissen und Kompetenzen von Lehrpersonen, etc.), so dass nicht nur aggregierte Individualmerkmale in die Analysen einbezogen werden. Auf Individualebene empfiehlt es sich zudem, die Arbeitsgedächtniskapazität sowie emotionale und motivationale Faktoren zu berücksichtigen, wobei Letztere auch als abhängige Variablen interessante Hinweise auf Fördereffekte bieten könnten. Hinsichtlich sprachlicher Kompetenzen sollte zudem eine direktere Erfassung über standardisierte Leistungstests in Erwägung gezogen werden. Da sich nebst Individualmerkmalen auch Klassenvariablen als bedeutsam für die mathematische Leistung erwiesen haben, spricht dies zudem auch zukünftig für statistische Analysemethoden, mit denen die hierarchische Datenstruktur im schulischen Kontext berücksichtigt werden kann. Hinsichtlich des Forschungsdesigns wäre des Weiteren die zeitliche Planung der ersten Datenerhebung und Datenauswertung für zukünftige Interventionsstudien zu überdenken. Dadurch könnte eine nach zentralen Merkmalen gewichtete Zuteilung der Klassen auf die Stichprobengruppen erfolgen, so dass sichergestellt werden kann, dass bedeutsame Merkmale in den Stichprobengruppen gleichmässig verteilt sind.

8. Literaturverzeichnis

Akremi, L., Baur, N. & Fromm, S. (2011). *Datenanalyse mit SPSS für Fortgeschrittene 1. Datenaufbereitung und uni- und bivariate Statistik* (3., überarbeitete und erweiterte Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Andersson, U. (2008a). Mathematical competencies in children with different types of learning difficulties. *Journal of Educational Psychology*, *100*(1), 48–66.

Andersson, U. (2008b). Working memory as a predictor of written arithmetical skills in children: The importance of central executive functions. *British Journal of Educational Psychology*, *78*(2), 181–203.

Andersson, U. (2010). Skill development in different components of arithmetic and basic cognitive functions: Findings from a 3-year longitudinal study of children with different types of learning difficulties. *Journal of Educational Psychology*, *102*(1), 115–134.

Andersson, U. & Östergren, R. (2012). Number magnitude processing and basic cognitive functions in children with mathematical learning disabilities. *Learning and Individual Differences*, *22*(6), 701–714.

Ansari, D., Garcia, N., Lucas, E., Hamon, K. & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport*, *16*(16), 1769–1773.

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review*, *2*(3), 213–236.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2016). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (14., überarbeitete und aktualisierte Aufl.). Berlin: Springer.

Baddeley, A. D. (1986). *Working memory*. Oxford: University Press.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49*(1), 5–28.

Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.

Baumert, J., Brunner, M., Lüdtke, O. & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? – Resultate kumulativer Wissenserwerbsprozesse. *Psychologische Rundschau*, *58*(2), 118–128.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.

Beck, E., Baer, M., Guldimann, T., Bischoff, S., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrerkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.

Bless, G., Schüpbach, M. & Bonvin, P. (2004). *Klassenwiederholung. Determinanten, Wirkungen und Konsequenzen*. Bern: Haupt.

Blum, W. & Leiss, D. (2005). Modellieren im Unterricht mit der "Tanken"-Aufgabe. *Mathematik lehren*, (128), 18–21.

Bochnik, K. & Ufer, S. (2016). Die Rolle (fach-)sprachlicher Kompetenzen zur Erklärung mathematischer Kompetenzunterschiede zwischen Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Grundschulforschung*, *9*(1), 133–147.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Zur Theorie sozialer Ungleichheiten* (S. 183–198). Göttingen: Schwarz.

Breitenbach, E. (2016). Intelligenz. In M. Dederich, I. Beck, U. Bleidick & G. Antor (Hrsg.), *Handlexikon der Behindertenpädagogik* (3., erweiterte und überarbeitete Aufl., S. 359–360). Stuttgart: Kohlhammer.

Brühwiler, C. & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, *21*(1), 95–108.

Brühwiler, C., Helmke, A. & Schrader, F.-W. (2017). Determinanten der Schulleistung. In M. K.W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (3., überarbeitete und aktualisierte Aufl., S. 291–314). Wiesbaden: Springer.

Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N. & Chavez, M. M. (2008). Mathematics intervention for first- and second-grade students with mathematics difficulties. The effects of Tier 2 intervention delivered as booster lessons. *Remedial and Special Education*, *29*(1), 20–32.

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Hughes Pfannenstiel, K., Porterfield, J. et al. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children*, *78*(1), 7–23.

Bundesamt für Statistik & Schweizerische Konferenz der kantonalen Erziehungsdirektoren. (2002). *Für das Leben gerüstet? Die Grundkompetenzen der Jugendlichen – Nationaler Bericht der Erhebung PISA 2000*. Neuenburg: Bundesamt für Statistik.

Bundesamt für Statistik & Schweizerische Konferenz der kantonalen Erziehungsdirektoren. (2004). *PISA 2003: Kompetenzen für die Zukunft. Erster nationaler Bericht*. Neuenburg: Bundesamt für Statistik.

Busch, J., Oranu, N., Schmidt, C. & Grube, D. (2013). Rechenschwäche im Grundschulalter: Reduzierte Verfügbarkeit basalen arithmetischen Faktenwissens und Belastung des Arbeitsgedächtnisses bei Drittklässlern. *Lernen und Lernstörungen*, *2*(4), 217–227.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, *46*(1), 3–18.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P. & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*(4), 499–531.

Case, R. & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, *61*(1–2), 27–58.

Cattell, R. B. (1965). *The scientific analysis of personality*. Harmondsworth: Penguin.

Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1: CFT 1. Test* (5., revidierte Aufl.). Göttingen: Hogrefe.

Cipolotti, L. & Butterworth, B. (1995). Toward a multiroute model of number processing: Impaired number transcoding with preserved calculation skills. *Journal of Experimental Psychology: General*, *124*(4), 375–390.

Codding, R. S., Burns, M. K. & Lukito, G. (2011). Meta-analysis of mathematic basis-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice*, *26,* 135–147.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Curran, P. J., West, S. G. & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29.

De Smedt, B. & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, *108*(2), 278–292.

De Smedt, B., Noël, M.-P., Gilmore, C. K. & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S. & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*(1), 83–120.

Desoete, A., Roeyers, H. & DeClercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities*, *37*(1), 50–61.

Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie. (2007). *Leitlinien zu Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter* (3., überarbeitete Aufl.). Köln: Deutscher Ärzte-Verlag.

Deutsche Gesellschaft für Psychologie. (2016). *Richtlinien zur Manuskriptgestaltung* (4., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.

Devine, A., Soltész, F., Nobes, A., Goswami, U. & Szűcs, D. (2013). Gender differences in developmental dyscalculia depend on diagnostic criteria. *Learning and Instruction*, *27,* 31–39.

Dilling, H. & Freyberger, H. J. (2014). *Taschenführer zur ICD-10-Klassifikation psychischer Störungen. Nach dem Pocket Guide von J. E. Cooper* (7., überarbeitete Aufl. unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2014). Bern: Huber.

Dilling, H., Mombour, W. & Schmidt, M. H. (2014). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F). Klinisch-diagnostische Leitlinien* (9. Aufl., unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2014). Bern: Huber.

Dirks, E., Spyer, G., van Lieshout, E. C. D. M. & de Sonneville, L. (2008). Prevalence of combined reading and arithmetic disabilities. *Journal of Learning Disabilities*, *41*(5), 460–473.

Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5., vollständig überarbeitete Aufl.). Berlin: Springer.

Dummert, F., Endlich, D., Schneider, W. & Schwenck, C. (2014). Entwicklung schriftsprachlicher und mathematischer Leistungen bei Kindern mit und ohne Migrationshintergrund. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *46*(3), 115–132.

Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, *60*(3), 163–183.

Ehlert, A., Fritz, A., Arndt, D. & Leutner, D. (2013). Arithmetische Basiskompetenzen von Schülerinnen und Schülern in den Klassen 5 bis 7 der Sekundarstufe. *Journal für Mathematik-Didaktik*, *34*(2), 237–263.

Ehlert, A., Schroeders, U. & Fritz-Stratmann, A. (2012). Kritik am Diskrepanzkriterium in der Diagnostik von Legasthenie und Dyskalkulie. *Lernen und Lernstörungen*, *1*(3), 169–184.

Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics. A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127.

Enders, C. K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138.

Ennemoser, M. & Krajewski, K. (2007). Effekte der Förderung des Teil-Ganzes-Verständnisses bei Erstklässlern mit schwachen Mathematikleistungen. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, *76*(3), 228–240.

Ennemoser, M., Krajewski, K. & Sinner, D. (2017). *MBK 1+. Test mathematischer Basiskompetenzen ab Schuleintritt*. Göttingen: Hogrefe.

Ennemoser, M., Sinner, D. & Krajewski, K. (2015). Kurz- und langfristige Effekte einer entwicklungsorientierten Mathematikförderung bei Erstklässlern mit drohender Rechenschwäche. *Lernen und Lernstörungen*, *4*(1), 43–59.

Esser, G., Wyschkon, A. & Ballaschk, K. (2008). *Basisdiagnostik umschriebener Entwicklungsstörungen im Grundschulalter*. Göttingen: Hogrefe.

Falkai, P. & Wittchen, H.-U. (2015). *Diagnostisches und statistisches Manual psychischer Störungen DSM-5®. Deutsche Ausgabe*. Göttingen: Hogrefe.

Feigenson, L., Carey, S. & Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, *44*(1), 33–66.

Feigenson, L., Dehaene, S. & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Fend, H. (1981). *Theorie der Schule* (2. Aufl.). München: Urban & Schwarzenberg.

Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Feuser, G. (1989). Allgemeine integrative Pädagogik und entwicklungslogische Didaktik. *Behindertenpädagogik*, *28*(1), 4–48.

Feuser, G. (2013). Die "Kooperation am Gemeinsamen Gegenstand". *Behinderte Menschen – Zeitschrift für gemeinsames Leben, Lernen und Arbeiten*, (2), 17–35.

Field, A. (2009). *Discovering statistics using SPSS* (3. Aufl.). London: SAGE. Retrieved from http://www.uk.sagepub.com/field3e/main.htm

Fischbach, A., Schuchardt, K., Brandenburg, J., Klesczewski, J., Balke-Melcher, C., Schmidt, C. et al. (2013). Prävalenz von Lernschwächen und Lernstörungen. Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen*, *2*(2), 65–76.

Fischer, C., Kopmann, H., Rott, D., Veber, M. & Zeinz, H. (2014). Adaptive Lehrerkompetenz und pädagogische Haltung. Lehrerbildung für eine inklusive Schule. *Jahrbuch für Allgemeine Didaktik*, *4,* 16–33.

Fischer, U., Roesch, S. & Moeller, K. (2017). Diagnostik und Förderung bei Rechenschwäche. *Lernen und Lernstörungen*, *6*(1), 25–38.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A. & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, *38*(2), 98–108.

Freesemann, O. (2014). *Schwache Rechnerinnen und Rechner fördern: Eine Interventionsstudie an Haupt-, Gesamt- und Förderschulen*. Wiesbaden: Springer.

Freudenthal, H. (1974). Die Stufen im Lernprozess und die heterogene Lerngruppe im Hinblick auf die Middenschool. *Neue Sammlung*, (14), 161–172.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H. & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10,* 29–44.

Fritz, A. & Ricken, G. (2008). *Rechenschwäche*. München: Reinhardt.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D. & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*(3), 493–513.

Fuchs, L. S. & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without comorbid reading disabilities. *Journal of Learning Disabilities*, *35*(6), 563–573.

Fuchs, L. S. & Fuchs, D. (2005). Enhancing mathematical problem solving for students with disabilities. *The Journal of Special Education*, *39*(1), 45–57.

Fuchs, L. S., Fuchs, D. & Compton, D. L. (2012). The early prevention of mathematics difficulty: Its power and limitations. *Journal of Learning Disabilities*, *45*(3), 257–269.

Fuchs, L. S., Fuchs, D. & Prentice, K. (2004). Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities*, *37*(4), 293–306.

Fuchs, L. S., Powell, S. R., Hamlett, C. L., Fuchs, D., Cirino, P. T. & Fletcher, J. M. (2008). Remediating computational deficits at third grade: A randomized field trial. *Journal of Research on Educational Effectiveness*, *1*(1), 2–32.

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer.

Gaidoschik, M. (2012). First-graders' development of calculation strategies: How deriving facts helps automatize facts. *Journal für Mathematik-Didaktik*, *33*(2), 287–315.

Gaidoschik, M. (2016). *Einmaleins verstehen, vernetzen, merken. Strategien gegen Lernschwierigkeiten* (3. Aufl.). Seelze: Klett/Kallmeyer.

Garrote, A., Sermier Dessemontet, R. & Moser Opitz, E. (2017). Facilitating the social participation of pupils with special educational needs in mainstream schools. A review of school-based interventions. *Educational Research Review*, *20,* 12–23.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological and genetic components. *Psychological Bulletin*, *114*(2), 345–362.

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K. & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development*, *24*(4), 411–429.

Geary, D. C., Bow-Thomas, C. & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, *54,* 372–391.

Geary, D. C., Hamson, C. O. & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, *77*(3), 236–263.

Geary, D. C., Hoard, M. K. & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning Disabilities*, *45*(4), 291–307.

Geary, D. C., Hoard, M. K., Byrd-Craven, J. & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88,* 121–151.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L. & Numtee, C. (2007). Cognitive mechanism underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*(4), 1343–1359.

Geary, D. C., Hoard, M. K. & Hamson, C. O. (1999). Numerical and arithmetical cognition. Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology*, *74,* 213–239.

Geary, D. C., Hoard, M. K., Nugent, L. & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*, *104*(1), 206–223.

Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Gelman, R. & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P. & Flojo, J. R. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202–1242.

Gersten, R., Jordan, N. C. & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 293–304.

Gerster, H.-D. (1996). Vom Fingerrechnen zum Kopfrechnen. In G. Eberle & R. Kornmann (Hrsg.), *Lernschwierigkeiten und Vermittlungsprobleme im Mathematikunterricht an Grund- und Sonderschulen. Möglichkeiten der Vermeidung und Überwindung* (S. 137–162). Weinheim.

Ginsburg, H. P. (1997). Mathematics learning disabilities: A view from developmental psychology. *Journal of Learning Disabilities*, *30*(1), 20–33.

Gniewosz, B. (2015). Experiment. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (2. Aufl., S. 83–91). Wiesbaden: Springer.

Gold, A. (2016). *Lernen leichter machen. Wie man im Unterricht mit Lernschwierigkeiten umgehen kann*. Göttingen: Vandenhoeck & Ruprecht.

Goldhaber, D. D. & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*(2), 129–145.

Goldman, S. R. (1989). Strategy instruction in mathematics. *Learning Disability Quarterly*, *12*(1), 43–55.

Greenberg, D. F. & Philips, J. A. (2013). Hierarchical linear modeling of growth curve trajectories using HLM. In G. D. Garson (Hrsg.), *Hierarchical linear modeling. guide and applications* (S. 219–247). Thousand Oaks: SAGE.

Gross-Tsur, V., Manor, O. & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology*, *38*(1), 25–33.

Grube, D. & Hasselhorn, M. (2006). Längsschnittliche Analysen zur Lese-, Rechtschreib- und Mathematikleistung im Grundschulalter: zur Rolle von Vorwissen, Intelligenz, phonologischem Arbeitsgedächtnis und phonologischer Bewusstheit. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven* (S. 87–105). Münster: Waxmann.

Grünke, M. (2006). Zur Effektivität von Fördermethoden bei Kindern und Jugendlichen mit Lernstörungen. *Kindheit und Entwicklung*, *15*(4), 239–254.

Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *Heidelberger Rechentest (HRT 1–4): Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.

Hanich, L. B., Jordan, N. C., Kaplan, D. & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, *93*(3), 615–626.

Hardesty, F. P. & Priester, H. J. (1966). *HAWIK. Hamburg-Wechsler-Intelligenztest für Kinder* (3. Aufl.). Bern: Verlag Hans Huber.

Häsel-Weide, U. (2016). *Vom Zählen zum Rechnen. Struktur-fokussierende Deutungen in kooperativen Lernumgebungen*. Wiesbaden: Springer.

Hasselhorn, M. & Gold, A. (2017). *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren* (4., aktualisierte Aufl.). Stuttgart: Kohlhammer.

Hasselhorn, M. & Schuchardt, K. (2006). Lernstörungen. Eine kritische Skizze zur Epidemiologie. *Kindheit und Entwicklung*, *15*(4), 208–215.

Hein, J., Bzufka, M. W. & Neumärker, K.-J. (2000). The specific disorder of arithmetic skills. Prevalence studies in a rural and an urban population sample and their clinic-neuropsychological validation. *European Child and Adolescent Psychiatry*, *9,* 87–101.

Helmke, A. (2010). Unterrichtsqualität. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4., überarbeitete und erweiterte Aufl., S. 886–895). Weinheim: Beltz.

Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 71–176). Göttingen: Hogrefe.

Hengartner, E. & Wieland, G. (2008). *Schweizer Zahlenbuch 3*. Zug: Klett & Balmer.

Hess, R. (2010). Retention. In C. S. Clauss-Ehlers (Hrsg.), *Encyclopedia of Cross-Cultural School Psychology* (S. 818–820). Boston, MA: Springer US.

Hill, H. C., Rowan, B. & Loewenberg Ball, D. (2005). Effects of teacher's mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371–406.

Höhtker, B. & Selter, C. (1995). Von der Hunderterkette zum leeren Zahlenstrahl. In G. N. Müller & E. C. Wittmann (Hrsg.), *Mit Kindern rechnen* (S. 122–137). Frankfurt: Arbeitskreis Grundschule.

Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2. Aufl.). New York: Routledge.

Huck, L. & Schröder, A. (2016). Psychosoziale Belastungen und Lernschwierigkeiten. *Lernen und Lernstörungen*, *5*(3), 157–164.

Hugener, I., Krammer, K. & Pauli, C. (2008). Kompetenzen der Lehrpersonen im Umgang mit Heterogenität: Differenzierungsmassnahmen im Mathematikunterricht. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise – Analyse und Bedeutung unterrichtlichen Handelns* (S. 47–66). Münster: Waxmann.

Hußmann, S. & Prediger, S. (2007). Mit Unterschieden rechnen – Differenzieren und Individualisieren. *Praxis der Mathematik in der Schule*, *49*(17), 1–8.

Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139–155.

Ise, E., Dolle, K., Pixner, S. & Schulte-Körne, G. (2012). Effektive Förderung rechenschwacher Kinder. Eine Metaanalyse. *Kindheit und Entwicklung*, *21*(3), 181–192.

Jacobs, C. & Petermann, F. (2003). Dyskalkulie – Forschungsstand und Perspektiven. *Kindheit und Entwicklung*, *12*(4), 197–211.

Jacobs, C. & Petermann, F. (2005). Diagnostik von Rechenstörungen. In M. Hasselhorn, H. Marx & W. Schneider (Hrsg.), *Diagnostik von Mathematikleistungen* (S. 71–104). Göttingen: Hogrefe.

Jahnke-Klein, S. (2001). *Sinnstiftender Mathematikunterricht für Mädchen und Jungen*. Baltmannsweiler: Schneider-Verlag.

Jiménez Gonzalez, J. E. & Garcia Espinel, A. I. (1999). Is IQ-achievement discrepancy relevant in the definition of arithmetic learning disabilities. *Learning Disability Quarterly*, *22,* 291–301.

Jiménez Gonzalez, J. E. & Garcia Espinel, A. I. (2002). Strategy choice in solving arithmetic word problems: Are there differences between students with learning disabilities, G-V poor performance and typical achievement students? *Learning Disability Quarterly*, *25,* 113–122.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, *30*(3), 420–437.

Jordan, N. C., Glutting, J. & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, *20*(2), 82–88.

Jordan, N. C. & Hanich, L. B. (2003). Characteristics of children with moderate mathematics deficiencies: A longitudinal perspective. *Learning Disabilities Research & Practice*, *18*(4), 213–221.

Jordan, N. C., Hanich, L. B. & Kaplan, D. (2003). A longitudinal study of mathematic competencies in children with specific mathematics difficulties versus children with comorbid mathematics and difficulties. *Child Development*, *74*(3), 834–850.

Jordan, N. C. & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews*, *15,* 60–68.

Kaufmann, L., Handl, P. & Thöny, B. (2003). Evaluation of a numeracy intervention program focusing on basic numerical knowledge and conceptual knowledge: A pilot study. *Journal of Learning Disabilities*, *36*(6), 564–573.

Kaufmann, L., Koppelstaetter, F., Siedentopf, C., Haala, I., Haberlandt, E., Zimmerhackl, L. B. et al. (2006). Neural correlates of the number-size interference task in children. *Neuroreport*, *17*(6), 587–591.

Kaufmann, L., Nuerk, H.-C., Graf, M., Krinzinger, H., Delazer, M. & Willmes, K. (2008). *Test zur Erfassung numerisch-rechnerischer Kompetenzen vom Kindergarten bis zur 3. Klasse (TEDI-MATH)*. Bern: Huber.

Kaufmann, L., Vogel, S. E., Starkey, M., Kremser, C. & Schocke, M. (2009). Numerical and non-numerical ordinality processing in children with and without developmental dyscalculia: Evidence from fMRI. *Cognitive Development*, *24,* 486–494.

Keller, B. & Noelle Müller, B. (2012). *Mathematik 3. Themenbuch Primarstufe*. Zürich: Lehrmittelverlag Zürich.

Kiel, E. (2017). Unterrichtsforschung im Kontext der empirischen Bildungsforschung. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Handbuch Bildungsforschung* (S. 1–22). Wiesbaden: Springer.

Koch, K. (2008). Förderung mathematischer Kompetenzen. In M. Fingerle & S. Ellinger (Hrsg.), *Sonderpädagogische Förderprogramme im Vergleich. Orientierungshilfen für die Praxis.* (S. 85–108). Stuttgart: Kohlhammer.

Kohn, J., Richtmann, V., Rauscher, L., Kucian, K., Käser, T., Grond, U. et al. (2013). Das Mathematikangstinterview (MAI). Erste psychometrische Gütekriterien. *Lernen und Lernstörungen*, *2*(3), 177–189.

Krajewski, K. (2003). *Vorhersage von Rechenschwäche in der Grundschule*. Hamburg: Verlag Dr. Kovač.

Krajewski, K. (2005). Vorschulische Mengenbewusstheit von Zahlen und ihre Bedeutung für die Früherkennung von Rechenschwäche. In M. Hasselhorn, H. Marx & W. Schneider (Hrsg.), *Diagnostik von Mathematikleistungen* (S. 49–70). Göttingen: Hogrefe.

Krajewski, K. (2007). Entwicklung und Förderung der vorschulischen Mengen-Zahlen-Kompetenz und ihre Bedeutung für die mathematischen Schulleistungen. In G. Schulte-Körne (Hrsg.), *Legasthenie und Dyskalkulie: Aktuelle Entwicklungen in Wissenschaft, Schule und Gesellschaft* (S. 325–332). Bochum: Winkler.

Krajewski, K. (2008a). Prävention der Rechenschwäche. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 360–370). Göttingen: Hogrefe.

Krajewski, K. (2008b). Vorschulische Förderung bei beeinträchtigter Entwicklung mathematischer Kompetenzen. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 122–135). Stuttgart: Kohlhammer.

Krajewski, K. & Ennemoser, M. (2013). Entwicklung und Diagnostik der Zahl-Grössen-Verknüpfung zwischen 3 und 8 Jahren. In M. Hasselhorn, A. Heinze & W. Schneider (Hrsg.), *Diagnostik mathematischer Kompetenzen* (S. 41–65). Göttingen: Hogrefe.

Krajewski, K., Küspert, P. & Schneider, W. (2002). *Deutscher Mathematiktest für erste Klassen: DEMAT 1+*. Göttingen: Beltz.

Krajewski, K., Nieding, G. & Schneider, W. (2007). *Mengen, zählen, Zahlen. Die Welt der Mathematik verstehen. Die große Förderbox*. Berlin: Cornelsen.

Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht*, *53*(4), 246–262.

Krajewski, K. & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*, *19*(6), 513–526.

Krammer, K., Lipowsky, F., Pauli, C., Schnetzler, C. L. & Reusser, K. (2012). Unterrichtsvideos als Medium zur Professionalisierung und als Instrument der Kompetenzerfassung von Lehrpersonen. In M. Kobarg, C. Fischer, I. M. Dalehefte, F. Trepke & M. Menk (Hrsg.), *Lehrerprofessionalisierung wissenschaftlich begleiten – Strategien und Methoden* (89-86). Münster: Waxmann.

Krauthausen, G. (1995). Die "Kraft der Fünf" und das denkende Rechnen. In G. N. Müller & E. C. Wittmann (Hrsg.), *Mit Kindern rechnen* (S. 87–108). Frankfurt/M: Arbeitskreis Grundschule.

Krauthausen, G. & Scherer, P. (2007). *Einführung in die Mathematikdidaktik* (3. Aufl.). Heidelberg: Spektrum Akademischer Verlag.

Krinzinger, H. & Günther, T. (2013). Lesen, Schreiben, Rechnen – gibt es Unterschiede zwischen den Geschlechtern? *Lernen und Lernstörungen*, *2*(1), 35–49.

Kroesbergen, E. H. & van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs. A meta-analysis. *Remedial and Special Education*, *24*(2), 97–114.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F. et al. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, *57*(3), 782–795.

Kucian, K., Loenneker, T., Dietrich, T., Dosch, M., Martin, E. & von Aster, M. (2006). Impaired neural networks for approximate calculation in dyscalculic children: A functional MRI study. *Behavioral and Brain Functions*, *2*(1), 31.

Kucian, K., Loenneker, T., Martin, E. & von Aster, M. (2011). Non-symbolic numerical distance effect in children with and without developmental dyscalculia: A parametric fMRI study. *Developmental Neuropsychology*, *36*(6), 741–762.

Kuhn, J.-T., Raddatz, J., Holling, H. & Dobel, C. (2013). Dyskalkulie vs. Rechenschwäche: Basisnumerische Verarbeitung in der Grundschule. *Lernen und Lernstörungen*, *2*(4), 229–247.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.

Lambert, K. (2015). *Rechenschwäche. Grundlagen, Diagnostik und Förderung*. Göttingen: Hogrefe.

Landerl, K., Bevan, A. & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8-9-year-old students. *Cognition*, *93*(2), 99–125.

Landerl, K. & Kaufmann, L. (2008). *Dyskalkulie. Modelle, Diagnostik, Intervention*. München: Reinhardt.

Langer, W. (2009). *Mehrebenenanalysen. Eine Einführung für Forschung und Praxis* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Langhorst, P., Hildenbrand, C., Ehlert, A., Ricken, G. & Fritz, A. (2013). Mathematische Bildung im Kindergarten – Evaluation des Förderprogramms "Mina und der Maulwurf" und Betrachtung von Fortbildungsvarianten. In M. Hasselhorn, A. Heinze & W. Schneider (Hrsg.), *Diagnostik mathematischer Kompetenzen* (S. 113–134). Göttingen: Hogrefe.

Lebens, M., Graff, M. & Mayer, P. (2011). The affective dimensions of mathematical difficulties in schoolchildren. *Education Research International*, (5), 1–13.

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D. et al. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767.

Lehmann, R. H. (2006). Zur Bedeutung der kognitiven Heterogenität von Schulklassen für den Lernstand am Ende der Klassenstufe 4. In A. Schründer-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung* (S. 109–121). Wiesbaden: VS Verlag für Sozialwissenschaften.

Leiss, D. (2010). Adaptive Lehrerinterventionen beim mathematischen Modellieren – empirische Befunde einer vergleichenden Labor- und Unterrichtsstudie. *Journal für Mathematik-Didaktik*, *31*(2), 197–226.

Leuders, T. & Prediger, S. (2016). *Flexibel differenzieren und fokussiert fördern im Mathematikunterricht*. Berlin: Cornelsen.

Lewis, C., Hitch, G. J. & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry*, *35*(2), 283–292.

Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz.

Lindberg, S. M., Hyde, J. S., Petersen, J. L. & Linn, M. C. (2010). New trends in gender and mathematics performance. A meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135.

Lipowsky, F. (2011). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 398–417). Münster: Waxmann.

Lorenz, J. H. & Radatz, H. (1993). *Handbuch des Förderns im Mathematikunterricht*. Hannover: Schroedel.

Lüdtke, O., Robitzsch, A., Trautwein, U. & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, *34*(2), 120–131.

Mabbott, D. J. & Bisanz, J. (2008). Computational skills, working memory, and conceptual knowledge in older children with mathematics learning disabilities. *Journal of Learning Disabilities*, *41*(1), 15–28.

Maehler, C. & Schuchardt, K. (2011). Working memory in children with learning disabilities. Rethinking the criterion of discrepancy. *International Journal of Disability, Development and Education*, *58*(1), 5–17.

Mähler, C. (2019). *Diskrepanzkriterium. In M. A. Wirtz (Hrsg.), Dorsch – Lexikon der Psychologie*. Verfügbar unter https://portal.hogrefe.com/dorsch/diskrepanzkriterium

Manger, T. & Eikeland, O. J. (1998). The effects of mathematical achievement and cognitive ability on girl's and boy's mathematics-selfconcept. *Zeitschrift für Pädagogische Psychologie*, *12*(4), 210–218.

Mazzocco, M. M. M. (2005). Challenges and identifying target skills for math disability screening and intervention. *Journal of Learning Disabilities*, *38*(4), 318–323.

Mazzocco, M. M. M., Devlin, K. T. & McKenney, S. J. (2008). Is it a fact? Timed arithmetic performance of children with mathematical learning disabilities (MLD) varies as a function of how MLD is defined. *Developmental Neuropsychology*, *33*(3), 318–344.

Mazzocco, M. M. M. & Räsänen, P. (2013). Contributions of longitudinal studies to evolving definitions and knowledge of developmental dyscalculia. *Trends in Neuroscience and Education*, *2*(2), 65–73.

McCloskey, M., Caramazza, A. & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196.

McGraw-Hill. (2002). *CTB terra nova*. Monterey, CA: CTB McGraw-Hill.

McLean, J. & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, *74*(3), 240–260.

Milz, I. (1993). *Rechenschwächen erkennen und behandeln. Teilleistungsstörungen im mathematischen Denken*. Dortmund.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L. & Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance - A longitudinal study on numerical development. *Research in Developmental Disabilities*, *32*(5), 1837–1851.

Montague, M. (2011). Effective instruction in mathematics for students with learning difficulties. In C. Wyatt-Smith, J. Elkins & S. Gunn (Hrsg.), *Multiple perspectives on difficulties in learning literacy and numeracy* (S. 295–313). Dordrecht: Springer.

Montague, M., Krawec, J. L., Enders, C. K. & Dietz, S. (2014). The effects of cognitive strategy instruction on math problem solving of middle-school students of varying ability. *Journal of Educational Psychology*, *106*(2), 469–481.

Moser Opitz, E. (2008a). Schulische Heilpädagogik. In F. Baier & S. Schnurr (Hrsg.), *Schulische und schulnahe Dienste. Angebote, Praxis und fachliche Perspektiven* (S. 57–86). Bern: Haupt.

Moser Opitz, E. (2008b). *Zählen, Zahlbegriff, Rechnen. Theoretische Grundlagen und eine empirische Untersuchung zum mathematischen Erstunterricht in Sonderklassen* (3. Aufl.). Bern: Haupt.

Moser Opitz, E. (2009). Integrativer Unterricht. Überlegungen zum Mathematiklernen. *Grundschule*, *41*(3), 12–15.

Moser Opitz, E. (2013). *Rechenschwäche/Dyskalkulie. Theoretische Klärungen und empirische Studien an betroffenen Schülerinnen und Schülern.* (2. Aufl.). Bern: Haupt.

Moser Opitz, E., Freesemann, O., Grob, U. & Prediger, S. (2016). *BASIS-MATH-G 4⁺–5. Gruppentest zur Basisdiagnostik Mathematik für das vierte Quartal der 4. Klasse und für die 5. Klasse*. Bern: Hogrefe.

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I. & Hußmann, S. (2017). Remediation for students with mathematics difficulties: An intervention study in middle schools. *Journal of Learning Disabilities*, *50*(6), 724–736.

Moser Opitz, E. & Ramseier, E. (2012). Rechenschwach oder nicht rechenschwach? Eine kritische Auseinandersetzung mit Diagnosekonzepten, Klassifikationssystemen und Diagnoseinstrumenten unter besonderer Berücksichtigung von älteren Schülerinnen und Schülern. *Lernen und Lernstörungen*, *1*(2), 99–117.

Moser Opitz, E. & Schindler, V. (2017). Mathematiklernen im Kontext von sprachlichen Faktoren. In A. Fritz, S. Schmidt & G. Ricken (Hrsg.), *Handbuch Rechenschwäche. Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie* (3., vollständig überarbeitete und erweiterte Aufl., S. 141–155). Weinheim: Beltz.

Moser Opitz, E., Stöckli, M., Grob, U., Nührenbörger, M. & Reusser, L. (im Druck). *BASIS-MATH-G 2⁺*. Gruppentest zur Basisdiagnostik Mathematik für das vierte Quartal der 2. Klasse und das erste Quartal der 3. Klasse. Bern: Hogrefe.

Moser Opitz, E., Stöckli, M., Grob, U., Nührenbörger, M. & Reusser, L. (2019). *BASIS-MATH-G 3⁺*. Gruppentest zur Basisdiagnostik Mathematik für das vierte Quartal der 3. Klasse und das erste Quartal der 4. Klasse. Bern: Hogrefe.

Müller, G. N. & Wittmann, E. C. (1984). *Der Mathematikunterricht in der Primarstufe*. Braunschweig: Vieweg.

Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B. & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, *40*(5), 458–478.

Mussolin, C., Mejias, S. & Noël, M.-P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition*, *115*(1), 10–25.

Muthén, B. (1997). Latent variable modeling of longitudinal and multilevel data. *Sociological Methodology*, *27*(1), 453–480.

Nezlek, J. B., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung. Vorteile und Möglichkeiten der Mehrebenenmodellierung mit Zufallskoeffizienten. *Psychologische Rundschau*, *57*(4), 213–223.

Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus*. Münster: Waxmann.

Noël, M.-P. & Rousselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers in Human Neuroscience*, *5,* 1–4.

Noël, M.-P. & Seron, X. (1993). Arabic number reading deficit: A single case study or when 236 is read (2306) and judged superior to 1258. *Cognitive Neuropsychology*, *10*(4), 317–339.

Nolte, M. (2015). Sprechen wir die gleiche Sprache? Kommentar zu Ennemoser, Sinner & Krajewski (2015). Kurz- und langfristige Effekte einer entwicklungsorientierten Mathematikförderung bei Erstklässlern mit drohender Rechenschwäche. Lernen und Lernstörungen, 4 (1) , 43-59. *Lernen und Lernstörungen*, *4*(1), 61–64.

Opdenakker, M.-C., van Damme, J., de Fraine, B., van Landeghem, G. & Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *School Effectiveness and School Improvement*, *13*(4), 399–427.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*(3), 345–357.

Paetsch, J., Felbrich, A. & Stanat, P. (2015). Der Zusammenhang von sprachlichen und mathematischen Kompetenzen bei Kindern mit Deutsch als Zweitsprache. *Zeitschrift für Pädagogische Psychologie*, *29*(1), 19–29.

Paetsch, J., Radmann, S., Felbrich, A., Lehmann, R. H. & Stanat, P. (2016). Sprachkompetenz als Prädiktor mathematischer Kompetenzentwicklung von Kindern deutscher und nichtdeutscher Familiensprache. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *48*(1), 27–41.

Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G. et al. (2018). Teachers' instructional adaptations: a research synthesis. *Review of Educational Research*, *88*(2), 205–242.

Paulus, C. (2009). *Die "Bücheraufgabe" zur Bestimmung des kulturellen Kapitals bei Grundschülern*. Verfügbar unter http://hdl.handle.net/20.500.11780/3344

Peng, P. & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties. Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities*, *49*(1), 3–20.

Pfister, M. (2016). *Adaptive Lernunterstützung im integrativen Mathematikunterricht: eine Videostudie*. Dissertation, Universität Zürich. Verfügbar unter http://opac.nebis.ch/ediss/20162747.pdf

Pfister, M., Moser Opitz, E. & Pauli, C. (2015). Scaffolding for mathematics teaching in inclusive primary classrooms: A video study. *ZDM Mathematics Education*, *47*(7), 1079–1092.

Piaget, J. & Szeminska, A. (1975). *Die Entwicklung des Zahlbegriffs beim Kinde*. Stuttgart: Klett.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D. et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41.

Pinel, P., Dehaene, S., Rivière, D. & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage*, *14*(5), 1013–1026.

Pinel, P., Le Clec, H. G., van de Moortele, P. F., Naccache, L., Le Bihan, D. & Dehaene, S. (1999). Event-related fMRI analysis of the cerebral circuit for number comparison. *Neuroreport*, *10*(7), 1473–1479.

Pixner, S. & Kaufmann, L. (2013). Prüfungsangst, Schulleistung und Lebensqualität bei Schülern. *Lernen und Lernstörungen*, *2*(2), 111–124.

Powell, S. R., Cirino, P. T. & Malone, A. S. (2017). Child-level predictors of responsiveness to evidence-based mathematics intervention. *Exceptional Children*, *83*(4), 359–377.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T. & Fletcher, J. M. (2009a). Do word-problem featured differentially affect problem difficulty as a function of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities*, *42*(2), 99–110.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T. & Fletcher, J. M. (2009b). Effects of fact retrieval tutoring on third-grade students with math difficulties with and without reading difficulties. *Learning Disabilities Research & Practice*, *24*(1), 1–11.

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E. & Benholz, C. (2015). Sprachkompetenz und Mathematikleistung – Empirische Untersuchung sprachlich bedingter Hürden in den Zentralen Prüfungen 10. *Journal für Mathematik-Didaktik*, *36*(1), 77–104.

Price, G. R., Holloway, I. D., Rasanen, P., Vesterinen, M. & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, *17*(24), R1042–1043.

Radatz, H., Schipper, W., Dröge, R. & Ebeling, A. (1996). *Handbuch für den Mathematikunterricht. 1. Schuljahr. Anregungen für die Unterrichtspraxis*. Hannover: Schroedel.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2. Aufl.). Thousand Oaks: SAGE.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T. & Du Toit, M. (2011). *HLM 7. Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International, Inc.

Raudenbush, S. W., Bryk, A. S. & Congdon, R. T. (2013). HLM 7.01 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

Rauscher, L., Kohn, J., Käser, T., Kucian, K., McCaskey, U., Wyschkon, A. et al. (2017). Effekte des Calcularis -Trainings. *Lernen und Lernstörungen*, *6*(2), 75–86.

Rauscher, L., Kohn, J., Käser, T., Mayer, V., Kucian, K., McCaskey, U. et al. (2016). Evaluation of a computer-based training program for enhancing arithmetic skills and spatial number representation in primary school children. *Frontiers in Psychology*, *7,* 913.

Reigosa-Crespo, V., Valdes-Sosa, M., Butterworth, B., Estevez, N., Rodriguez, M., Santos, E. et al. (2012). Basic numerical capacities and prevalence of developmental dyscalculia:. The Havana Survey. *Developmental Psychology*, *48*(1), 123–135.

Resnick, L. B. (1983). A developmental theory of number understanding. In H. P. Ginsburg (Hrsg.), *The development of mathematical thinking* (S. 109–151). New York: Academic Press.

Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist*, *44*(2), 162–169.

Ricken, G., Fritz, A. & Balzer, L. (2011). Mathematik und Rechnen – Test zur Erfassung von Konzepten im Vorschulalter (MARKO-D) – ein Beispiel für einen niveauorientierten Ansatz. *Empirische Sonderpädagogik*, *3*(3), 256–271.

Ricken, G., Fritz-Stratmann, A. & Balzer, L. (2013). *MARKO-D. Mathematik- und Rechenkonzepte im Vorschulalter – Diagnose*. Göttingen: Hogrefe.

Robinson-Cimpian, J. P., Theule Lubienski, S., Ganley, C. M. & Copur-Gencturk, Y. (2014). Teachers' perceptions of students' mathematics proficiency may exacerbate early gender gaps in achievement. *Developmental Psychology*, *50*(4), 1262–1281.

Rogalla, M. & Vogt, F. (2008). Förderung adaptiver Lehrerkompetenz: eine Interventionsstudie. *Unterrichtswissenschaft*, *36*(1), 7–36.

Saß, H., Wittchen, H.-U., Zaudig, M. & Houben, I. (2003). *Diagnostisches und statistisches Manual psychischer Störungen – Textversion – DSM-IV-TR*. Göttingen: Hogrefe.

Schäfer, J. (2005). *Rechenschwäche in der Eingangsstufe der Hauptschule. Lernstand, Einstellungen und Wahrnehmungsleistungen. Eine empirische Studie*. Hamburg: Verlag Dr. Kovač.

Scherer, P. (1995). *Entdeckendes Lernen im Mathematikunterricht der Schule für Lernbehinderte. Theoretische Grundlegung und evaluierte unterrichtspraktische Erprobung*. Heidelberg: Schindele.

Scherer, P. (2009). Produktives Mathematiklernen – auch in Förderschule?! In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche – Lernwege, Schwierigkeiten, Hilfen* (2., erweiterte und aktualisierte Aufl., S. 434–447). Weinheim: Beltz.

Scherer, P. & Moser Opitz, E. (2010). *Fördern im Mathematikunterricht der Primarstufe*. Heidelberg: Spektrum Akademischer Verlag.

Schipper, W. (2002). Thesen und Empfehlungen zum schulischen und ausserschulischen Umgang mit Rechenstörungen. *Journal für Mathematik-Didaktik*, *23*(3/4), 243–261.

Schmassmann, M. & Moser Opitz, E. (2008a). *Heilpädagogischer Kommentar zum Schweizer Zahlenbuch 2. Hinweise zum Arbeiten mit Kindern mit mathematischen Lernschwierigkeiten* (2. vollständig überarbeitete Aufl.). Zug: Klett & Balmer.

Schmassmann, M. & Moser Opitz, E. (2008b). *Heilpädagogischer Kommentar zum Schweizer Zahlenbuch 3. Hinweise zum Arbeiten mit Kindern mit mathematischen Lernschwierigkeiten*. Zug: Klett & Balmer.

Schneider, H. (2009). Nachweis und Behandlung von Multikollinearität. In S. Albers, D. Klapper, U. Konradt, A. Walter & J. Wolf (Hrsg.), *Methodik der empirischen Forschung* (3., überarbeitete und erweiterte Aufl., S. 221–236). Wiesbaden: Gabler Verlag.

Schneider, W., Küspert, P. & Krajewski, K. (2013). *Die Entwicklung mathematischer Kompetenzen*. Paderborn: Ferdinand Schöningh.

Schuchardt, K., Brandenburg, J., Fischbach, A., Büttner, G., Grube, D., Mähler, C. et al. (2015). Die Entwicklung des akademischen Selbstkonzeptes bei Grundschulkindern mit Lernschwierigkeiten. *Zeitschrift für Erziehungswissenschaft*, *18*(3), 513–526.

Schuchardt, K., Kunze, J., Grube, D. & Hasselhorn, M. (2006). Arbeitsgedächtnisdefizite bei Kindern mit schwachen Rechen- und Schriftsprachleistungen. *Zeitschrift für Pädagogische Psychologie*, *20*(4), 261–268.

Schuchardt, K. & Mähler, C. (2010). Unterscheiden sich Subgruppen rechengestörter Kinder in ihrer Arbeitsgedächtniskapazität, im basalen arithmetischen Faktenwissen und in den numerischen Basiskompetenzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*(4), 217–225.

Schuchardt, K., Piekny, J., Grube, D. & Mähler, C. (2014). Einfluss kognitiver Merkmale und häuslicher Umgebung auf die Entwicklung numerischer Kompetenzen im Vorschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *46*(1), 24–34.

Schulte-Körne, G. (2014). Spezifische Lernstörungen. Vom DSM-IV zum DSM-5 [Specific learning disabilities - from DSM-IV to DSM-5]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *42*(5), 369-374.

Schulte-Körne, G., Deimel, W. & Remschmidt, H. (2001). Zur Diagnostik der Lese-Rechtschreibstörung [Diagnosis of reading and spelling disorder]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *29*(2), 113–116.

Schwenck, C. & Schneider, W. (2003). Einflussfaktoren für den Zusammenhang von Rechen- und Schriftsprachleistungen im frühen Grundschulalter. *Kindheit und Entwicklung*, *12*(4), 212–221.

Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. *Zeitschrift für Pädagogik*, *60*(6), 850–866.

Selter, C. (2000). Vorgehensweisen von Grundschüler(inne)n bei Aufgaben der Addition und Subtraktion im Zahlenraum bis 1000. *Journal für Mathematik-Didaktik*, *21*(2), 227–258.

Semper, I., Mende, L. & Berkemeyer, N. (2017). Schul- und Unterrichtsforschung. In T. Burger & N. Miceli (Hrsg.), *Empirische Forschung im Kontext Schule* (S. 31–48). Wiesbaden: Springer.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–22.

Sinner, D. (2011). *Prävention von Rechenschwäche durch ein Training mathematischer Basiskompetenzen in der ersten Klasse*. Dissertation, Universität Giessen. Verfügbar unter http://geb.uni-giessen.de/geb/volltexte/2011/8198/pdf/SinnerDaniel_2011_05_25.pdf

Snijders, T. A. B. & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, *22*(3), 342–363.

Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel analysis. An introduction to basic and advanced multilevel modeling* (2. Aufl.). London: SAGE.

Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P. et al. (2017). Math anxiety and its relationship with basic arithmetic skills among primary school children. *The British Journal of Educational Psychology*, *87*(3), 309–327.

Stanat, P. (2006). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* (S. 189–219). Wiesbaden: VS Verlag für Sozialwissenschaften.

Staub, F. C. (2014). Fachunterrichtscoaching auf der Grundlage des Content-Focused Coaching. In K. Mattern & U. Hirt (Hrsg.), *Coaching im Fachunterricht. Wie Unterrichtsentwicklung gelingt* (S. 39–52). Weinheim: Beltz.

Stern, E. (2009). Früh übt sich: Neuere Ergebnisse aus der LOGIK-Studie zum Lösen mathematischer Textaufgaben. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche – Lernwege, Schwierigkeiten, Hilfen* (2., erweiterte und aktualisierte Aufl., S. 151–164). Weinheim: Beltz.

Stöckli, M., Moser Opitz, E., Pfister, M. & Reusser, L. (2014). Gezielt fördern, differenzieren und trotzdem gemeinsam lernen. Überlegungen zum inklusiven Mathematikunterricht. *Sonderpädagogische Förderung heute*, *59*(1), 44–56.

Stubbe, T. C., Tarelli, I. & Wendt, H. (2012). Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 231–246). Münster: Waxmann.

Swanson, H. L. & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, *76*(2), 249–274.

Swanson, H. L., Orosco, M. J. & Lussier, C. M. (2014). The effects of mathematics strategy instruction for children with serious problem-solving difficulties. *Exceptional Children*, *80*(2), 149–168.

Szűcs, D., Devine, A., Soltész, F., Nobes, A. & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*(10), 2674–2688.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5. Aufl.). Boston: Pearson / Allyn and Bacon.

Tarelli, I., Schwippert, K. & Stubbe, T. C. (2012). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 247–267). Münster: Waxmann.

Thomas, K., Schulte-Körne, G. & Hasselhorn, M. (2015). Stichwort – Entwicklungsstörungen schulischer Fertigkeiten. *Zeitschrift für Erziehungswissenschaft*, *18*(3), 431–451.

Thompson, S. K. (2002). *Sampling* (2. Aufl.). New York, NY: Wiley. Retrieved from http://www.loc.gov/catdir/description/wiley036/2001046957.html

Tiedemann, J. & Billmann-Mahecha, E. (2004). Kontextfaktoren der Schulleistung im Grundschulalter. Ergebnisse aus der Hannoverschen Grundschulstudie. *Zeitschrift für Pädagogische Psychologie*, *18*(2), 113–124.

Toll, S. W. M., Kroesbergen, E. H. & van Luit, J. E. H. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *The British Journal of Educational Psychology*, *86*(3), 429–445.

Toll, S. W. M., van der Ven, S. H. G., Kroesbergen, E. H. & van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, *44*(6), 521–532.

Trautwein, U., Niggli, A., Schnyder, I. & Lüdtke, O. (2009). Between-teacher differences in homework assignments and the development of students' homework effort, homework emotions, and achievement. *Journal of Educational Psychology*, *101*(1), 176–189.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Hrsg.), *But what does it mean? The use of effect sizes in educational research* (S. 55–66). London: National Foundation for Educational Research.

University of Texas System & Texas Education Agency. (2007). *Texas early mathematics inventories-progress monitoring (TEMI-PM)*. Austin: Authors.

Van de Pol, J., Volman, M. & Beishuizen, J. (2010). Scaffolding in teacher–student interaction. A decade of research. *Educational Psychology Review*, *22*(3), 271–296.

Van de Walle, J. A. (2007). *Elementary and middle school mathematics. Teaching developmentally*. Boston: Allyn & Bacon.

Vanbist, K., Ghesquière, P. & de Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities*, *35*(11), 3001–3013.

Vogel, S. E. & Ansari, D. (2012). Neurokognitive Grundlagen der typischen und atypischen Zahlenverarbeitung. *Lernen und Lernstörungen*, *1*(2), 135–149.

Von Aster, M. (2005). Wie kommen Zahlen in den Kopf? Ein Modell der normalen und abweichenden Entwicklung zahlenverarbeitender Hirnfunktionen. In M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern: Neurowissenschaft, Psychologie, Pädagogik* (S. 13–33). Göttingen: Vandenhoeck & Ruprecht.

Von Aster, M., Kucian, K. & Martin, E. (2006). Gehirnentwicklung und Dyskalkulie. *Sprache, Stimme, Gehör*, *30*(4), 1–6.

Von Aster, M., Kucian, K., Schweiter, M. & Martin, E. (2005). Rechenstörungen im Kindesalter. *Monatszeitschrift für Kinderheilkunde*, *153*(7), 614–622.

Von Aster, M., Schweiter, M. & Weinhold Zulauf, M. (2007). Rechenstörungen bei Kindern. Vorläufer, Prävalenz und psychische Symptome. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *39*(2), 85–96.

Von Aster, M., Weinhold Zulauf, M. & Horn, R. (2006). *ZAREKI-R – Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern*. Frankfurt am Main: Harcourt.

Vorstand des Arbeitskreises Schweiz-Liechtenstein der GDM. (2013). Planarbeit als Qualitätsfalle für den Mathematikunterricht. Eine Stellungnahme aus mathematikdidaktischer Perspektive. *GDM-Mitteilungen*, *94,* 21–22. Verfügbar unter http://didaktik-der-mathematik.de/pdf/gdm-mitteilungen-94.pdf

Voss, T., Kunter, M., Seiz, J., Hoehne, V. & Baumert, J. (2014). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. *Zeitschrift für Pädagogik*, *60*(2), 184–201.

Vukovic, R. K. & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology*, *115*(2), 227–244.

Vukovic, R. K. & Siegel, L. S. (2010). Academic and cognitive characteristics of persistent mathematics difficulty from first through fourth grade. *Learning Disabilities Research & Practice*, *25*(1), 25–38.

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Cambridge: Harvard University Press.

Waldis, M., Grob, U., Pauli, C. & Reusser, K. (2010). Der Einfluss der Unterrichtsgestaltung auf Fachinteresse und Mathematikleistung. In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 209–251). Münster: Waxmann.

Wartha, S. & Schulz, A. (2011). *Aufbau von Grundvorstellungen (nicht nur) bei besonderen Schwierigkeiten im Rechnen. Handreichungen des Programms SINUS an Grundschulen*. Verfügbar unter http://www.sinus-an-grundschulen.de/fileadmin/uploads/Material_aus_SGS/Handreichung_WarthaSchulz.pdf

Wechsler, D. (1964). *Die Messung der Intelligenz Erwachsener* (3. Aufl.). Bern: Huber.

Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1: CFT 1. Handanweisung* (5. revidierte Aufl.). Göttingen: Hogrefe.

Wember, F. B. (2009). Mathematik unterrichten – eine subsidiäre Aktivität? Nicht nur bei Kindern mit Lernschwierigkeiten! In P. Scherer (Hrsg.), *Produktives Lernen für Kinder mit Lernschwächen: Fördern durch Fordern* (S. 239–245). Hamburg: Persen.

Werner, B. (2009). *Dyskalkulie – Rechenschwierigkeiten. Diagnose und Förderung rechenschwacher Kinder an Grund- und Sonderschulen*. Stuttgart: Kohlhammer.

Wiener, J. & Tardif, C. Y. (2004). Social and emotional functioning of children with learning disabilities: Does special education placement make a difference? *Learning Disabilities Research & Practice*, *19*(1), 20–32.

Wißmann, J., Heine, A., Handl, P. & Jacobs, A. M. (2013). Förderung von Kindern mit isolierter Rechenschwäche und kombinierter Rechen- und Leseschwäche: Evaluation eines numerischen Förderprogramms für Grundschüler. *Lernen und Lernstörungen*, *2*(2), 91–109.

Wittmann, E. C. (1992). Üben im Lernprozess. In E. C. Wittmann & G. N. Müller (Hrsg.), *Handbuch produktiver Rechenübungen. Band 2. Vom halbschriftlichen zum schriftlichen Rechnen.* (S. 175–182). Stuttgart/Düsseldorf.

Wittmann, E. C. & Müller, G. N. (Hrsg.). (1992). *Handbuch produktiver Rechenübungen. Band 2. Vom halbschriftlichen zum schriftlichen Rechnen.* Stuttgart/Düsseldorf.

Wocken, H. (1998). Gemeinsame Lernsituationen. Eine Skizze zur Theorie des gemeinsamen Unterrichts. In A. Hildeschmidt & I. Schnell (Hrsg.), *Integrationspädagogik. Auf dem Weg zu einer Schule für alle* (S. 37–52). Weinheim: Juventus.

Wolf, C. & Best, H. (2010). Lineare Regressionsanalyse. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 607–638). Wiesbaden: VS Verlag für Sozialwissenschaften.

Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100.

Woodward, J. & Brown, C. (2006). Meeting the curricular needs of academically low-achieving students in middle grade mathematics. *The Journal of Special Education*, *40*(3), 151–159.

Wu, S. S., Meyer, M. L., Maeda, U., Salimpoor, V., Tomiyama, S., Geary, D. C. et al. (2008). Standardized assessment of strategy use and working memory in early mental arithmetic performance. *Developmental Neuropsychology*, *33*(3), 365–393.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193.

Wyschkon, A., Kohn, J., Ballaschk, K. & Esser, G. (2009). Sind Rechenstörungen genau so häufig wie Lese-Rechtschreibstörungen? [Is a specific disorder of arithmetic skills as common as reading/spelling disorder?]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *37*(6), 499-512.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B. & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects students achievement* (REL Issues & Answers 033). Washinton D.C.: U.S. Departement of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Verfügbar unter https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf

Zurbriggen, C. (2016). *Schulklasseneffekte. Schülerinnen und Schüler zwischen komparativen und normativen Einflüssen*. Wiesbaden: Springer.

9. Abbildungsverzeichnis

10. Tabellenverzeichnis